PRTN #247

F. Tobagi
UCLA
January 17, 1978

On the Performance Analysis of Multihop Packet Radio Systems: Part II - Star Configuration Employing Slotted ALOHA*

1. Tree-Structured Networks: Definitions

A number of papers have appeared which study various simple network topologies. Single-hop networks, where terminals communicate directly with a central station have been investigated extensively; access modes in such environments have been carefully evaluated [2,3,4]. A two-hop configuration involving a ring of repeaters around a station have been analyzed by Gitman [5]: network capacity was derived, but no consideration was made of network delays. In this note, we consider centralized networks characterized by an array of repeaters organized in a symmetric tree configuration with a (single) station at the root; all devices are provided with omnidirectional antennas and employ slotted ALOHA random access over a single shared channel. (See Figure 1.) Traffic originates at terminals located at the outskirts of the tree, and is destined to the station. Thus, we consider inbound traffic. The first transmission of a packet from a terminal is to a repeater located at a leaf of the tree. The routing of the packet through the network is completely specified by the tree structure, as is the connectivity pattern among the devices. The basic performance measure sought is the throughput-delay trade-off and its dependence on such key system parameters as the topology (degree of the tree, depth of the tree,** and the connectivity

---

*The results discussed in this note have already been published in [1], and distributed as PRTN #234. This note, however, is more complete in that all derivations are included.

**In this note, however, we shall restrict ourselves to trees of depth 1.

pattern among devices), repeaters' retransmission delays and repeaters' storage capacities.

2. Traffic Model and Transmission Protocol

The time axis is assumed to be universal and slotted into segments whose duration is equal to the transmission time of a packet.* All devices are assumed to be synchronized and start packet transmission at the beginning of a slot. Associated with each repeater located at a leaf of the tree is a population of terminals which generate new packets at an aggregate rate of s packets per slot, all destined to the station. Each repeater is provided with a finite storage capacity which can accomodate a maximum of M packets. The station has an infinite storage capacity. Packets are transmitted by repeaters on a first-come-first-served basis; when its buffer is non-empty, a repeater transmits the head of its queue with a probability p. When the packet transport is successful (i.e., the transmission is free of interference and storage is available at the receiving repeater), the packet is deleted from the sender's queue; otherwise, the packet incurs a retransmission delay geometrically distributed, with mean 1/p. A repeater learns about its success or failure instantaneously; that is, acknowledgments are assumed to be instantaneous and for free. At any one time, a repeater can be either transmitting or receiving, but not both simultaneously. The station always has its receiver on. The packet processing time at any device is considered to be negligible.

Let R denote a repeater located at some level k in the tree. Let S(R) denote the set of repeaters which are sons of R. Let F(R) denote the father of R. A packet successfully transmitted by the population of repeaters S(R)

---

*Packets in this study are all of a fixed size.

can be "blocked" at the immediate destination repeater R; blocking is due to two factors: (i) either R or F(R)* or both are in a transmit mode or (ii) R's receiver is on, F(R) is quiet, but R's buffer is full.

Due to the blocking of traffic at the receiving repeater, the rate of successful transmissions of packets to a leaf from its corresponding population of terminals is actually greater than s and is denoted by λ. Furthermore, this process of packet arrivals to a leaf is assumed to be a Bernoulli one.

In this note, we shall consider the special case of trees of depth 1 (and degree N), referred to as star networks (See Fig. 2), and we shall provide an analysis by means of which the throughput-delay performance can be numerically evaluated. In a forthcoming note [6], we shall consider configurations in which all repeaters are in line-of-sight and within range of each other. These are referred to as fully-connected networks.

3. Analysis of Star Networks Employing Slotted ALOHA

Let N denote the degree of the tree. (N also denotes the number of repeaters present in the configuration.) Given the transmission protocol adopted (as described in section 2 above), and given that the input process to each repeater is a Bernoulli process, the state of the system at slot t is entirely defined by the vector

$$\underline{n}^t = (n_1^t, n_2^t, \ldots, n_N^t)$$

where n_i^t is the number of packets present at the i^{th} repeater at slot t.

---

*Given that the hearing pattern among repeaters is assumed to be determined by the tree structure, a transmission by F(R) over a slot interferes with a transmission from S(R) over the same slot.

Moreover, $\underline{n}^t$ is a Markov chain. In this section, we shall analyze $\underline{n}^t$ by first considering the case where the buffer size at each repeater is limited to one packet (M=1) and then by treating the more general case M>1. In all cases, the configurations considered will be symmetric.

3.1 The Single-Buffer Case

3.1.1 Analysis

In this symmetric single-buffer configuration, the state of the system can be equivalently described by the number of repeaters with non-empty buffers, referred to as the number of "active" repeaters. Let n^t $(0 \leq n^t \leq N)$ denote that number at slot t. The Markov chain n^t has a transition matrix P whose $(i,j)^{th}$ element is given by

$$p_{ij} = \begin{cases} 0 & j < i - 1 \\ P_s(i)(1 - \lambda)^{N-i} & j = i - 1 \\ [1 - P_s(i)] \binom{N - i}{k} \lambda^k (1 - \lambda)^{N-i-k} & \\ \quad + P_s(i) \binom{N - i}{k + 1} \lambda^{k+1} (1 - \lambda)^{N-i-(k+1)} & j = i + k \\ & k = 0, 1, 2, \ldots, N - i \end{cases} \tag{1}$$

where $P_s(i)$ denotes the probability of a successful transmission given i active repeaters and is expressed as

$$P_s(i) = ip(1 - p)^{i-1} \tag{2}$$

Let $\pi_i \overset{\Delta}{=} \lim_{t \to \infty} Pr\{n^t = i\}$. We compute the stationary distribution $\Pi = \{\pi_0, \pi_2, \ldots, \pi_N\}$ by solving recursively the system $\Pi = \Pi P$.

Let $\bar{n}$ denote the average number of active repeaters. We have

$$\bar{n} = \sum_{k=0}^{N} k\pi_k \tag{3}$$

A packet successfully transmitted by a population of terminals can be "blocked" at the immediate destination repeater. As already mentioned in section 2, blocking is due to two factors: (i) the repeater is in a transmit mode* (and we denote by β the probability of such an event) or (ii) the repeater's receiver is on but its packet buffer is full (and we denote by α the probability of this event). Let $B = \alpha + \beta$. We have:

$$\alpha = \Pr\{\text{queue at receiving repeater full}\}\ (1 - p) \tag{4}$$

$$= (1 - p)\bar{n}/N$$

$$\beta = [1 - \Pr\{\text{queue at receiving repeater empty}\}]p \tag{5}$$

$$= p\bar{n}/N$$

$$B = \bar{n}/N \tag{6}$$

The total network throughput, denoted by S, is defined as the rate of successful packets received at the station; it is given by

$$S = (N - \bar{n})\lambda \tag{7}$$

The packet delay D is defined to be the time since the packet is originated at the terminal until it is successfully received at the station. We distinguish two components: (i) the access delay D_a, defined to be the time required for the packet to be correctly received at the leaf repeater, and (ii) the network delay D_n which consists of the time elapsed since the packet is accepted at the leaf repeater until it is successsully received at the

---

*Recall that the station's receiver is assumed to be always on.

station. By Little's result, the average network delay is given by

$$D_n = \bar{n}/S \qquad (8)$$

The maximum value of λ, allowable in this model, is a function of the access mode in use by the terminals. Indeed, if a slotted ALOHA mode is used, it is well known that the maximum rate of successful packets that can be transmitted by an infinite population of terminals is $\lambda = 1/e = 0.368$. On the other hand, given the memoryless property of the Bernoulli input process, the above analysis corresponds also to the "linear-feedback" model whereby, following the successful transmission of its buffer, a repeater is assumed to generate a new packet after a geometrically distributed time with mean $1/\lambda$. In the linear-feedback model, the rate λ can take any value between 0 and 1. $B = \bar{n}/N$ represents the fraction of time a repeater is active; and D_n represents the total packet delay.

3.2 Numerical Results

The main difficulty in obtaining a closed form solution to the present problem lies in the fact that the outcome of the system in a slot (i.e., the amount of work completed) is dependent on the state of the system in that slot. Thus, the time spent by a packet until its success (i.e., the service time of the packet) is dependent on the evolution of the system over the period of time that constitutes its service. This situation is usually observed in the study of random-access techniques and renders the analysis of systems utilizing these techniques a rather tedious task. The numerical approach adopted here has helped overcome this difficulty (to a certain extent); its advantages over simulation are many: it is simpler, it is more economical, and the numerical results it provides are more accurate (if not exact). Accordingly it allows the generation of numerical results for a large range of system parameters. This is indeed a quality of great importance when one attempts to gain insight on the behavior of a system via experimental

design and it is reflected in the number of cases studied and presented in this note. The technique, however, does present a major drawback: it amounts to performing measurements on a model of the real system which is based on a number of simplifying assumptions (acknowledgments are instantaneous and for free, processing time is negligible...). The relaxation of such assumptions may cause the analysis to become intractable; while it is more readily performed in a simulation model. Nevertheless, the assumptions made are not unreasonable when one's objective is, again, to gain insight into the behavior of packet radio systems and to determine the trends followed by the performance measures as a function of the system parameters.

Fixing N and λ, we observe that $\bar{n}$ is a concave function of p. Thus there exists a value of p which minimizes $\bar{n}$. From Equations (6), (7) and (8) which express D_n, B and S in terms of $\bar{n}$ we note that D_n and B are also concave functions of p while S is a convex function of p. Moreover, it is clear that the value of p which maximizes S, minimizes D_n and B. As an example we show in Figure 3 D_n, S and B versus p for N = 3 and two values of λ. We observe that the throughput S is not as sensitive to p as are D_n and B. That is, if p is improperly tuned, while the system can maintain the throughput desired, the network delay D_n and the probability of blocking B (and thus the access delay) may suffer large increases!

In Figure 4, we plot the optimum delay versus the achieved throughput for various values of N. We note that as the degree of the tree increases, so does the network delay. The reverse behavior is observed for the probability of blocking B over a large range of S ($0 < S < 0.35$) as shown in Figure 5. The total packet delay D will be dealt with in a subsequent section. Next, we examine the network performance at saturation, i.e., at maximum

attainable throughput.

3.1.3 Network Saturation

In terms of the feedback model, network behavior at saturation is obtained by driving λ to 1. With $\lambda = 1$, the Markov chain n^t reduces to one with only two states (state N-1 and state N) as depicted in Figure 6. Thus, at saturation, the stationary distribution Π is such that $\pi_0 = \pi_1 = \ldots = \pi_{N-2} = 0$, and such that π_{N-1} and π_N are solution of the system

$$\pi_{N-1}\,(N-1)p(1-p)^{N-2} + \pi_N\,Np(1-p)^{N-1} = \pi_{N-1}$$
$$\pi_{N-1} + \pi_N = 1 \qquad (9)$$

The latter yields

$$\pi_{N-1} = \frac{Np(1-p)^{N-1}}{1 - (N-1)p(1-p)^{N-2} + Np(1-p)^{N-1}} \qquad (10)$$

π_{N-1} represents also the network throughput. Indeed, the average number of active repeaters $\bar{n}$ is given by

$$\bar{n} = (N-1)\pi_{N-1} + N\pi_N$$
$$= N - \pi_{N-1} \qquad (11)$$

Given $\lambda = 1$, we see from Eq. (7) that S_{out} is precisely π_{N-1}. The network delay D_n is given by

$$D_n = \frac{N}{\pi_{N-1}} - 1$$
$$= \frac{1}{p(1-p)^{N-1}} - \frac{(N-1)p}{1-p} \qquad (12)$$

Let us examine the special case of N = 2. The throughput and delay are expressed, as long as $p \neq 1$, as

$$S = \frac{2p}{1 + 2p} \tag{13}$$

$$D_n = 1 + \frac{1}{p} \tag{14}$$

The network capacity is obtained for $p \to 1$ and equals

$$\lim_{P \to 1} S = \frac{2}{3} \tag{15}$$

The minimum packet delay is then equal to 2 (see Fig. 4).

Let us now study the limiting case as $N \to \infty$. It is clear from Eq. (10) that, as N gets large, S is on the order of $Np(1-p)^{N-1}$. The maximum throughput is then achieved by assigning to p the value 1/N, in which case it is expressed as

$$S = (1 - \frac{1}{N})^{N-1} \tag{16}$$

This, in the limit as $N \to \infty$, approaches 1/e, which, as expected, is the slotted ALOHA channel capacity with infinite population. For $p = \frac{1}{N}$ the network delay is given by

$$D_n = \frac{\quad}{(1 - \frac{1}{N})^{N-1}} - 1 \tag{17}$$

which, as N gets large, follows the linear asymplotic behavior

$$D_n = Ne - 1 \tag{18}$$

The above results are summarized in Fig. 7 where we plot network capacity, network delay and the probability of blocking at saturation versus N.

Consider now the two-hop environment. In this case the maximum achievable value for λ is 1/e. The system capacity is then expressed as

$$\max_{p} \left\{ \frac{N}{e} \left[1 - B\left(\frac{1}{e}, p\right) \right] \right\}$$

In Fig. 8 we plot the optimum network throughput S (maximized over p, λ kept constant) versus λ for various values of N. The system capacity is precisely the network throughput at $\lambda = 1/e$. We note the following. For the larger values of N ($N > 3$), the throughput S (which increases with increasing values of λ) levels off rather rapidly and gets really close to its maximum value for $\lambda < 1/e$. This is not so with the smaller values of N; for N = 2 and 3, values of λ greater than 1/e can produce system throughput noticeably higher than what can be achieved at $\lambda = 1/e$. This leads us to state that, for all practical purposes, the limiting hop is the terminal-to-repeater hop for N = 2 and 3, and the repeater-to-station hop for larger N. Moreover, we note that the system capacity for N = 2 is smaller than for N = 3, and that for $N \geq 3$, it is a decreasing function of N. In Fig. 9 we plot the system capacity versus N for the two-hop configuration.

3.1.4 The Throughput-Delay Tradeoff

To complete the delay analysis, we need to evaluate the access delay D_a for a given throughput S. Let us first examine the various states a terminal can be in and the possible transitions that exist among the states. Fig. 10 represents the state diagram for the population of terminals associated with a repeater. First, a terminal is in the thinking state. After a random period of time, the terminal generates and transmits a new packet. If the transmission is unsuccessful due to a collision with other contending terminals the terminal joins the set of colliding terminals and reschedules transmission

of its packet following a random retransmission delay, which we denote here by X. The terminal retransmits its packet and repeats this process until its transmission is free of collision by other terminals. When the latter is true, then the packet will be successfully received at the repeater if and only if the repeater is not transmitting and its buffer is not full*; otherwise, the terminal joins the set of blocked terminals and reschedules transmission of its packet following the random retransmission delay. The process will repeat itself until the collision-free transmission of the packet is successfully received at the repeater, in which case the terminal rejoins the set of thinking terminals. It is clear from the diagram in Fig. 10 that the average access delay D_a is equal to the average time spent by a terminal in transiting from point A_1 to point A_5, and which we denote by $T_{A_1A_5}$. Assuming that the blocking probability B is uniform over time and independent of the state that the population of terminals is in**, we can write (referring again to the diagram in Fig. 10),

$$D_a = T_{A_1A_5} = T_{A_1A_2} + \left(\frac{1}{1-B} - 1\right)(T_{A_3A_4} + T_{A_4A_2}) \qquad (19)$$

To estimate $T_{A_1A_2}$ and $T_{A_3A_4}$, we call upon previously published results. Slotted ALOHA with an infinite population has been thoroughly analyzed by Kleinrock and Lam [7]. With the channel input from the infinite population

---

*In the case M = 1, this condition is equivalent to the repeater's buffer being empty.

**The analysis of these multi-hop environments has been simplified by the basic assumption that the processes governing the inner-hop (repeater-to-station) are stochastically independent from those governing the terminal-to-repeater hop; in particular we assume that the process of offered input to a repeater is independent from the state of the repeater (active or empty) and vice versa.

modeled as an independent Poisson process with an average of λ packets/slot, letting the maximum retransmission delay be an integer number K of packet slots (the retransmission delay being uniformly distributed over the K slots*), and neglecting the propagation delay, the delay $D_{S\text{-}ALOHA}(\lambda)$ is then given by [7], [8]

$$D_{S\text{-}ALOHA}(\lambda) = \left[1 + \frac{E(K+1)}{2}\right] \tag{20}$$

where

$$E = \frac{1-q_n}{q_t}$$

$$q_n = \left[e^{-\frac{G}{k}} + \frac{G}{k} e^{-G}\right]^k e^{-\lambda}$$

$$q_t = \frac{e^{-\frac{G}{k}} - e^{-G}}{1 - e^{-G}} \left[e^{-\frac{G}{k}} + \frac{G}{k} e^{-G}\right]^{k-1} e^{-\lambda}$$

$$\lambda = G \frac{q_t}{q_t + 1 - q_n}$$

The normalized delay (in slots) is shown in Fig. 11 versus λ for various values of K. For each value of λ, we note that an optimum value of K, K_{opt}, can be selected so as to achieve minimum delay. The lower envelope of all delay cuves provides the throughput-delay performance of slotted ALOHA with infinite population. Fig. 12 shows K_{opt} versus λ.

---

*The geometric retransmission randomization can be used as well. However, it was shown [8] that the slotted ALOHA channel performance is dependent primarily upon the average value of the retransmission delay and quite insensitive to its exact distribution.

The model used by Kleinrock and Lam in determing packet delay can be represented by the diagram of Fig. 13, where the delay $D_{S\text{-}ALOHA}(\lambda)$ corresponds to $T_{A_1'A_2'}$. Comparing the two diagrams in question, we note the following Both include representation of slotted ALOHA channels with throughput λ. However while in the diagram of Fig. 13 the entire input to the channel is produced by the independent Poisson process with rate λ, in the diagram of Fig. 10, only the fraction s is produced by an indep[illegible]t process; the remaining fraction $(\lambda - s)$ is generated by the terminals in the blocked state. Regarding transmissions by these terminals as "new" input to the channel, and neglecting the effect of the correlation in traffic introduced here by the rescheduling protocol, we approximate $T_{A_1A_2}$ and $T_{A_4A_2}$ by

$$T_{A_1A_2} = T_{A_4A_2} = D_{S\text{-}ALOHA}(\lambda) \tag{21}$$

$T_{A_3A_4}$ is the average rescheduling delay and is simply given by

$$T_{A_3A_4} = \frac{K_{opt}(\lambda)}{2} \tag{22}$$

and thus an estimate of D_a is given by

$$D_a = \frac{1}{1 - B} D_{S\text{-}ALOHA}(\lambda) + \frac{B}{1 - B} \frac{K_{opt}}{2} \tag{23}$$

In Fig. 14 we plot the packet delay $D_a + D_n$ versus S for various values of N. We basically note that for $N \geq 3$, the throughput-delay performance degrades slightly as N increases. As for the N = 2 throughput-delay curve, we note the following. First, the system capacity for N = 2 is, as indicated above, roughly equal to the one achieved with N = 5. Secondly, the curve exhibits the lowest delay for very small S, but as S increases to reach its

maximum value, the curve intercepts both the N = 3 and N = 5 delay curves. To explain this behavior, we note that for a given throughput S, D_n is smaller for smaller N (see Fig. 4) while D_a is larger, due to the fact that for smaller N a larger value of λ is required to achieve that throughput, of course provided that the latter does not exceed the system capacity (see Fig. 8).

3.1.5 Dynamic Control for Improved Performance

Assuming that each repeater knows exactly the state of the system in each slot*, then one can improve the performance by maximizing the instantaneous throughput, given by $P_s(i)$ in Eq. (2), with respect to the transmission parameter p. Given that $n^t = i$, $P_s(i)$ is maximized for $p = 1/i$.

We illustrate the gain obtained via the dynamic control by plotting in Figures 15 and 16 the network delay D_n and the blocking probability B respectively. We also plot, in Fig. 17, the system capacity versus N for the two-hop configuration, and in Fig. 18, the throughput-delay curves for various values of N. One point is worth noting here. For N > 3, the system capacity is attained for optimum values of λ smaller than 1/e; as λ

*In practical situations, the assumption that each repeater knows the exact current state of the system clearly does not hold. The repeaters have no means of communicating among themselves other than the channel itself. However each repeater may individually estimate the system state by observing the channel outcome over some period of time, and apply a control action based upon the estimate. In the context of a one-hop slotted ALOHA environment, Lam and Kleinrock [9] give some heuristic control-estimation algorithms which prove to be very satisfactory. The results obtained here assuming full knowledge of the system state will then represent the ultimate performance; a bound on the performance obtained via heuristic estimation algorithms.

increases beyond its optimum value the system throughput decreases, as shown in Fig. 19, and reaches the corresponding system capacity in the uncontrolled case. This explains the dashed portions of the N = 5 and N = 10 delay curves in Fig. 18.

3.2 The Multi-Buffer Case

3.2.1 Analysis

With M > 1, the state of the system is described by the vector $\underline{n}^t = (n_1^t , n_2^t , \ldots, n_N^t)$ where n_i^t $(0 \leq n_i^t \leq M)$ denotes the number of packets buffered at repeater i at time t. Let S denote the state space; we have

$$S = \{(n_1 , n_2 , \ldots, n_N) / 0 \leq n_i \leq M , \forall i = 1, 2, .., N\}$$

We are seeking the probability of the one-step transition from state $\underline{m} = (m_1 , m_2 , \ldots, m_N)$ to state $\underline{n} = (n_1 , n_2 , \ldots, n_N)$, which we denote by $\Pr [\underline{n} / \underline{m}]$. We distinguish the following cases.

First, in any slot, the amplitude of change in n_i^t cannot exceed one, and there can be at most one departure. Therefore we have

(i) If there exists an index i such that $|m_i - n_i| > 1$ or if there exist i, j, $i \neq j$, such that $n_i = m_i - 1$ and $n_j = m_j - 1$ then $\Pr \{\underline{n} / \underline{m}\} = 0$

If the above conditions are not satisfied, then either a successful transmission took place, or no packet was successfully transmitted. Thus, we have

(ii) If there exists an index i_0 such that $m_{i_0} = n_{i_0} + 1$ (indicating a successful transmission by repeater i_0) then

$$\Pr \{\underline{n} / \underline{m}\} = p \prod_{j \neq i_0} (1 - p)^{x_j} \prod_{\ell \neq i_0} [\lambda \xi_\ell^- + (1 - \lambda + \lambda \zeta_\ell) \zeta_\ell] \qquad (24$$

where

$$\chi_j = \begin{cases} 1 & \text{if } m_j > 0 \\ 0 & \text{if } m_j = 0 \end{cases}$$

$$\xi_j^- = \begin{cases} 1 & \text{if } m_j = n_j - 1 \\ 0 & \text{if } m_j \geq n_j \end{cases}$$

$$\xi_j = \begin{cases} 1 & \text{if } m_j = n_j \\ 0 & \text{if } m_j \neq n_j \end{cases}$$

$$\zeta_j = \begin{cases} 1 & \text{if } n_j = M \\ 0 & \text{if } n_j < M \end{cases}$$

The term $p \prod_{j \neq i_0} (1 - p)^{\chi_j}$ represents the probability that i_0 is the only transmitting repeater among all active ones. The second product term represents the probability of all changes (that is, the presence or absence of arrivals) which occured at the remaining queues in the current slot. The indicator ζ_ℓ accounts for the queue lengths being finite; an arrival which finds the queue full is rejected.

(iii) Otherwise no successful transmission took place. Letting $I_s = \{j \mid m_j = n_j\}$ we have

$$\Pr\{\underline{n} / \underline{m}\} = \prod_{j \in I_s} [p\chi_j + (1 - p)^{\chi_j}(1 - \lambda + \lambda\zeta_j)]$$

$$- \sum_{j \in I_s} p\chi_j \prod_{\substack{k \in I_s \\ k \neq j}} (1 - p)^{\chi_k}(1 - \lambda + \lambda\zeta_k) \prod_{j \notin I_s} (1 - p)^{\chi_j}\lambda \qquad (25)$$

where χ_j and ζ_j are as defined in (ii) above.

According to the model under consideration, an arrival to a repeater in a slot t is rejected (blocked) if that repeater is in

transmit mode during that slot. Thus, the number of packets queued at repeater $j \in I_s$ remains unchanged with probability $px_j + (1 - p)^{x_j} (1 - \lambda + \lambda\tau_j)$, provided that any transmission (represented by the term px_j) is unsuccessful. Since the repeaters are all independent, the probability of the event $\{m_j = n_j\}$ for all $j \in I_s$ is then given by the expression in the first bracket, in which the summation

$$\left[\sum_{j \in I_s} px_j \prod_{\substack{k \in I_s \\ k \neq j}} (1 - p)^{x_k} (1 - \lambda + \lambda\tau_k) \right]$$

represents the probability of all possible successful transmissions. Now, for all $j \notin I_s$, the number of packets increased by one; the probability of this event is simply given by the last product term.

The transition matrix P is computed numerically. Let $\Pi = \{\pi_{\underline{n}}\}_{\underline{n} \in S}$ be the stationary distribution of $\underline{n}^t$. Π is evaluated by iteratively solving the system $\Pi = \Pi P$. Let q_i denote the queue length at repeater i. The marginal distribution of q_i is given by

$$\Pr\{q_i = k\} = \sum_{\{\underline{n} \in S \mid n_i = k\}} \pi_{\underline{n}} \tag{26}$$

The average queue length is then given by

$$\bar{q} = \sum_{k=0}^{M} k \Pr\{q_i = k\} \tag{27}$$

The blocking probabilities α and β are expressed as

$$\alpha = Pr\{q = M\}\ (1 - p) \tag{28}$$

$$\beta = [1 - Pr\{q = o\}]\ p \tag{29}$$

The network throughput (in this symmetric case) is given by

$$S = N\ \lambda(1 - \alpha - \beta) \tag{30}$$

and by Little's result, the network delay is computed by

$$D_n = \frac{\bar{q}}{\lambda(1 - \alpha - \beta)} \tag{31}$$

3.2.3 Numerical Results

In Fig. 20 we plot on the (S,D_n) plane the constant λ contours (varying p) for N=3, M=2. The optimum delay is obtained by taking the lower envelope. It is noted that given λ, the value of p yielding optimum delay does not exactly correspond to the value of p which yields minimum blocking (and therefore maximum throughput). However, the probability of blocking at minimum delay is not significantly different from the minimum blocking achievable! The effect M has on network delay is shown in Fig. 21 where we plot, for N=2 and 3, the optimum delay curves corresponding to various values of M. The increase with larger M is due to the additional queueing time incurred. We also note a slight decrease in network capacity. The effect M has on the probability of blocking is shown in Fig. 22 where we plot the minimum blocking as a function of S. Note the (slight) decrease achieved by going from M=1 to M=2. M=3, however, offers no further significant improvement!

Thus, for a given network throughput S, an increase in M results in an increase in D_n and a decrease in D_a (due to a decrease in B). What is then

the effect on the total delay D? In Figures 23 and 24, we plot D versus S for N=2 and N=3, respectively, and various values of M. Again, we only note a slight improvement in performance by going from M=1 to M=2. No further significant improvement is gained beyond M=2. The increase in network capacity (observed particularly with N=2) is obviously due to the decrease in B.

The lack of important improvement experienced by increasing M is mainly explained by the fact that the system, at optimum is mostly "channel bound" as opposed to "storage bound". To show that, we consider the (α,β) plane on which we plot the constant λ contours. When p is small, α predominates: $\alpha>\beta$; As p increases, the inequality reverses. The locus of optima is displayed in Fig. 25 for various values of N. The curves corresponding to N=2 and N=3 lie almost entirely in the $\beta>\alpha$ half of the quadrant, showing that blocking is mostly due to the receiver being shut off. However, as N increases, the optimum drifts to the $\alpha>\beta$ region. This effect is due to the fact that, for the same throughput, the optimum p decreases as N increases in order to prevent conflict among a larger number of contending users. Is the system then storage bound when N is large, say 10 for example? It can be argued that there is still no significant improvement by increasing M. First, with large N, D_n is the predominant delay factor; indeed for a given S, D_n increases with N (see Fig. 3) while D_a decreases as does S/N (for N=10, S/N <0.04). Secondly, as S remains lower than 0.35 (value close to the capacity of these networks with large N), B is smaller for larger N rendering it ineffectual to further decrease it in an attempt to decrease D_a. For example, consider N=10 and S=0.35; we have: D_n=10, B=0.38 and D_a=2.5 yielding D=12.5. By taking B=0, we can decrease D_a to 1.15 providing thus a lower bound on D of 11.15, a rather unimportant improvement. Moreover, due to the queueing effect, D_n increases with larger M.

4. Conclusion

The above analysis has provided a means by which the throughput and delay performance measures are evaluated for slotted ALOHA packet radio star-networks. In particular, we have shown the effect on the system performance of various system parameters, namely the transmission probability p, the number of repeaters N and the repeater's buffer size M. The results lead us to believe that the system is channel bound (the processing time at repeaters assumed negligible); a slight improvement may be gained by increasing the buffer size to M=2, but no significant improvement is obtained beyond that. Moreover we considered a dynamically controlled transmission protocol which provides a tremendous improvement in system performance.

In two forthcoming notes [6,10] we shall report on the results of a similar study performed on fully-connected networks. Both slotted ALOHA and CSMA will be considered.

References

[1] F. Tobagi, "Performance Analysis of Packet Radio Communications Systems," Proceedings of the IEEE National Telecommunications Conference, December 1977, pp. 12:6-1 12:6-7.

[2] Kleinrock, L. and F. Tobagi, "Packet Switching in Radio Channels: Part I -Carrier Sense Multiple Access Modes and their Throughput Delay Characteristics," IEEE Transactions on Communications, vol. COM-23, No. 12, December 1975, pp. 1400-1416.

[3] Tobagi, F. and L. Kleinrock, "Packet Switching in Radio Channels: Part II - The Hidden Terminal Problem in Carrier Sense Multiple Access and the Busy Tone Solution," IEEE Transactions on Communications, Vol. COM-23, Vol. 12, December 1975, pp. 1417-1433.

[4] Tobagi, F.A. and L. Kleinrock, "Packet Switching in Radio Channels: Part III - Polling and (Dynamic) Split-channel Reservation Multiple Access," IEEE Transactions on Communications, Vol. COM-24, No. 8, August 1976, pp. 832-845.

[5] Gitman, I., "On the Capacity of Slotted ALOHA Networks and Some Design Problems," IEEE Transactions on Communications, Vol. COM-23 No. 3, March 1975, pp. 305-317.

[6] F. Tobagi, "On the Performance Analysis of Multihop Packet Radio Systems: Part III - Fully Connected Configurations Employing Slotted ALOHA," Packet Radio Temporary Note #248, SRI International Menlo Park, California,

[7] Kleinrock, L.and S. Lam, "Packet Switching in a Multiaccess Broadcast Channel: Performance Evaluation," IEEE Transactions on Communications, Vol. COM-23, pp. 410-423, April 1970.

[8] Lam, S., "Packet Switching in a Multiaccess Broadcast Channel with Applications to Satellite Communication in a Computer Network," School of Engineering and Applied Science, University of California, Los Angeles, UCLA-Eng-7429. April 1974 (also published as Ph.D. Dissertation).

[9] Lam, S. and L. Kleinrock, "Packet Switching in a Multiaccess Broadcast Channel: Dynamic Control Procedures," IEEE Transactions on Communications, Vol. COM-23, pp. 891-904, September 1975.

[10] Tobagi, F., "On the Performance Analysis of Multihop Packet Radio Systems: Part IV - Fully Connected Configurations Employing CSMA," Packet Radio Temporary Note #249, SRI International, Menlo Park, California.

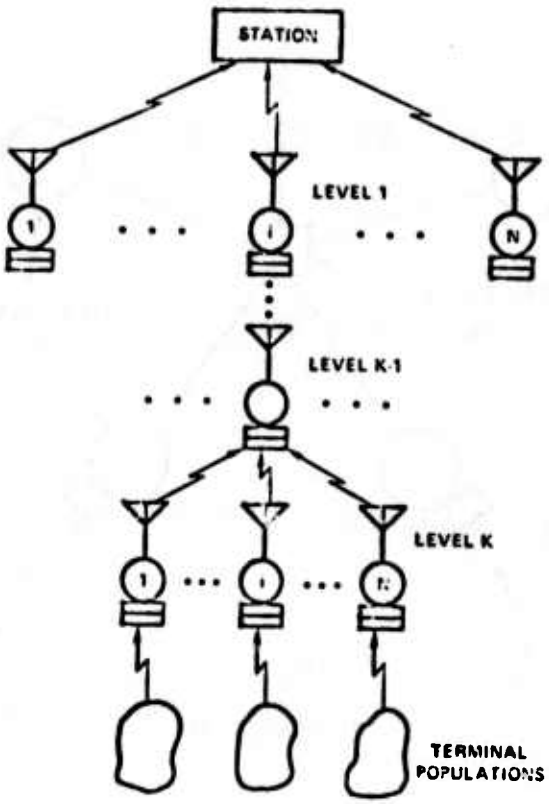

Figure 1 Tree-Structured Networks

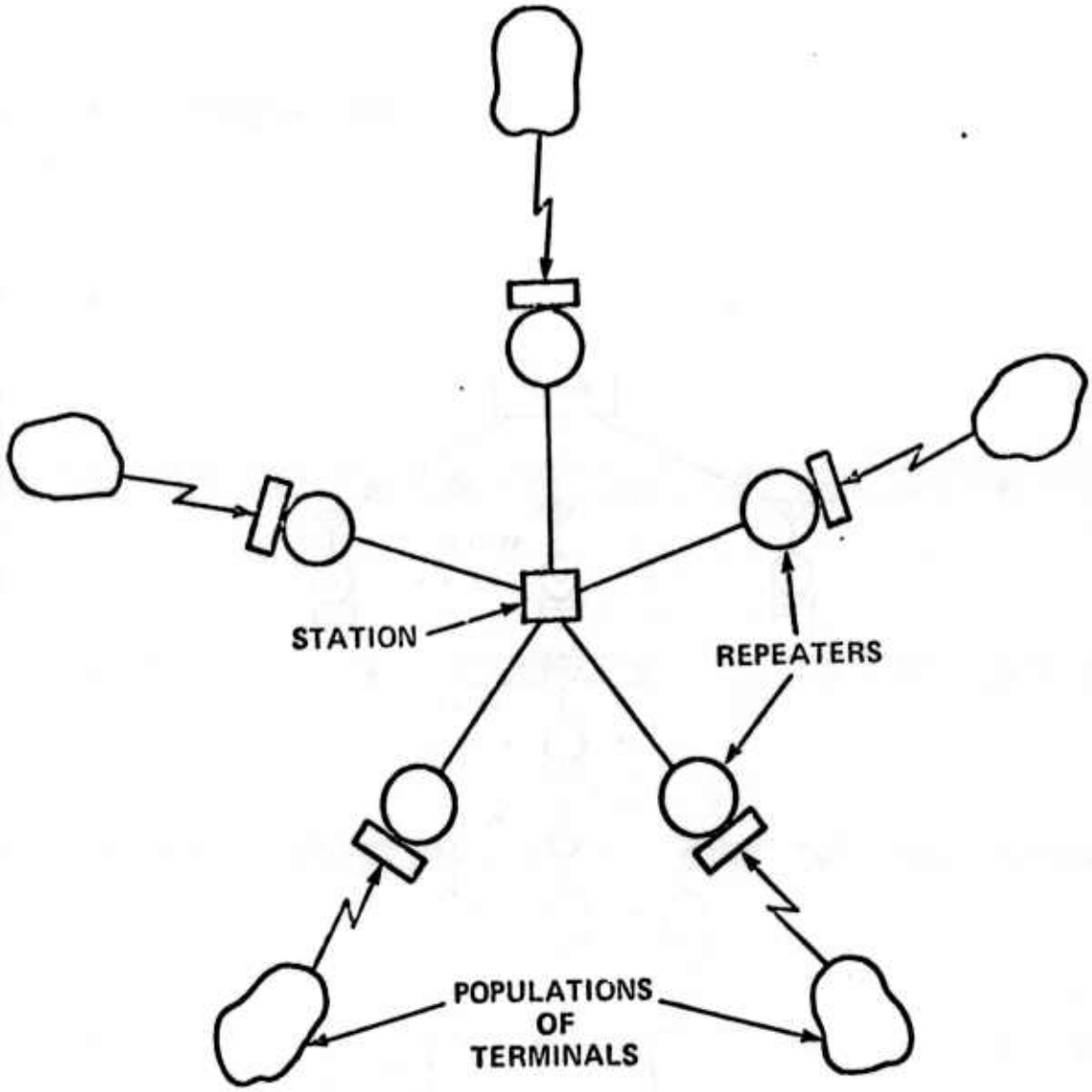

Figure 2 A Two-Hop Star Configuration

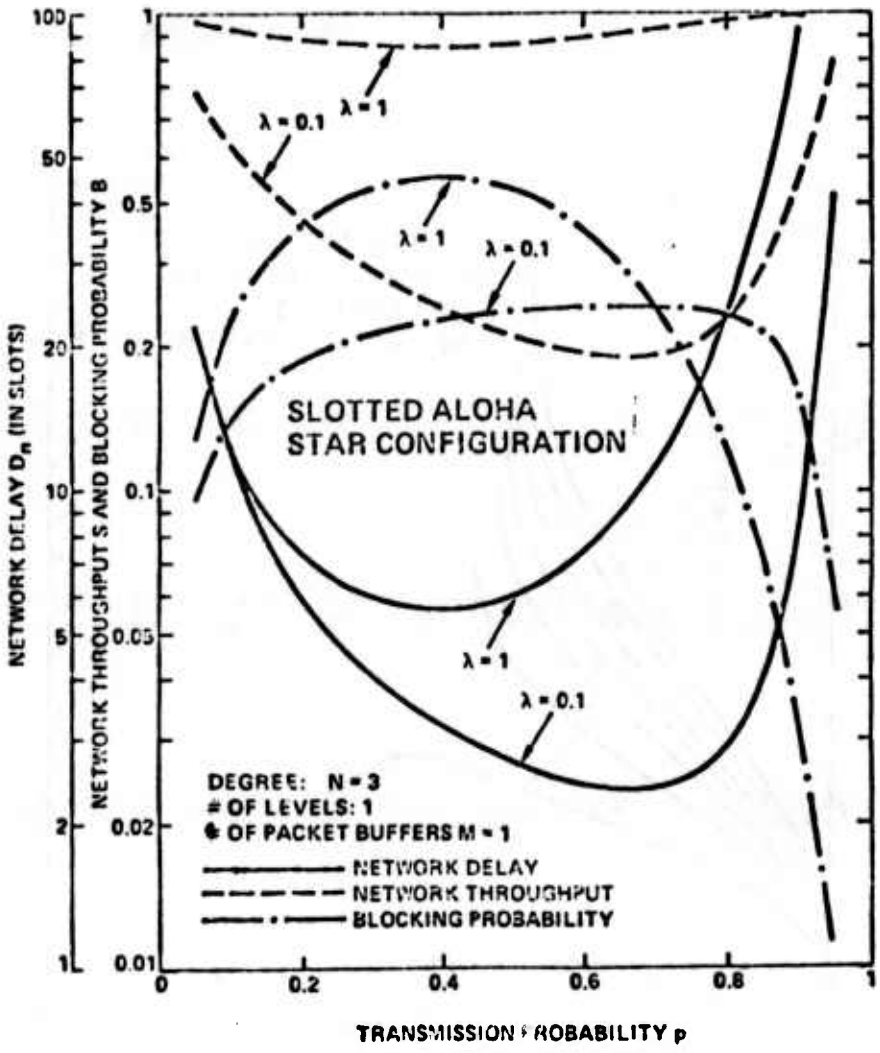

Figure 3 Slotted ALOHA Star Configuration: Network Delay, Throughput and Probability of Blocking versus p.

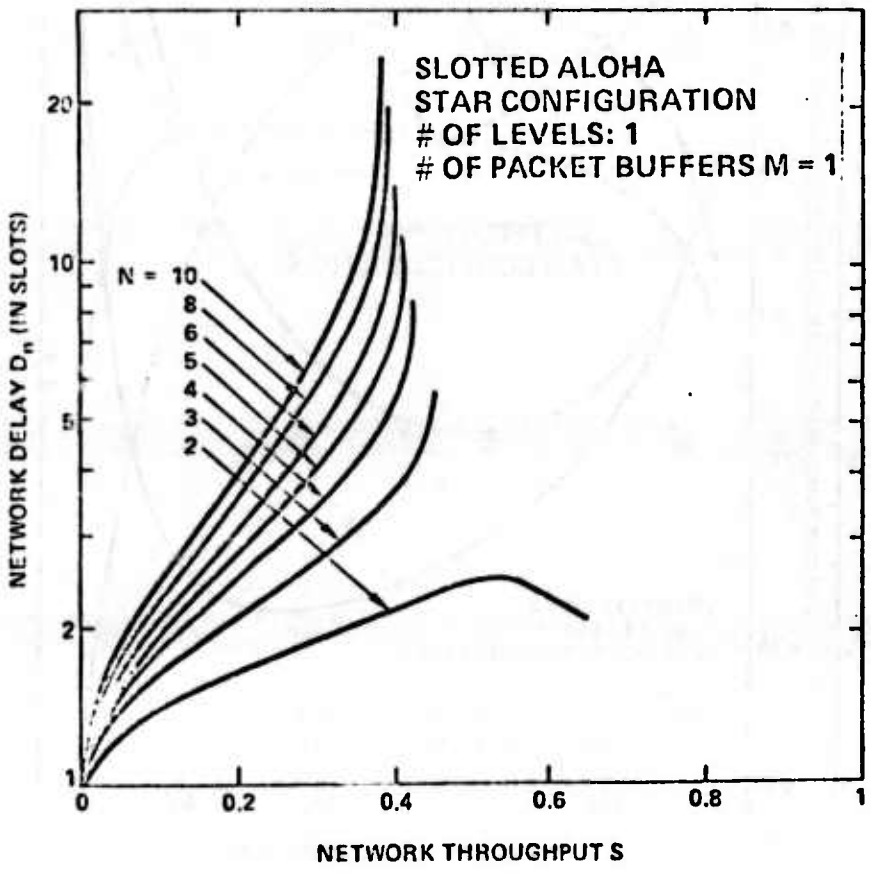

Figure 4 Slotted ALOHA Star Configuration: Optimum (Network) Throughput Delay Curves.

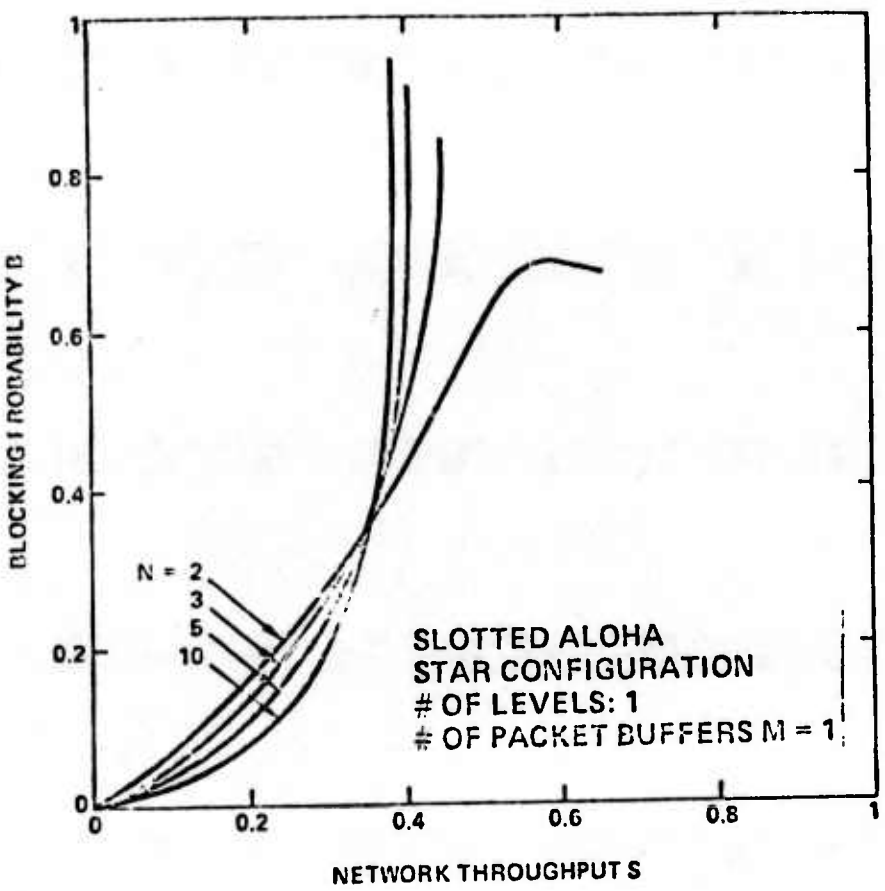

Figure 5 Slotted ALOHA Star Configuration: Minimum Blocking versus Throughput.

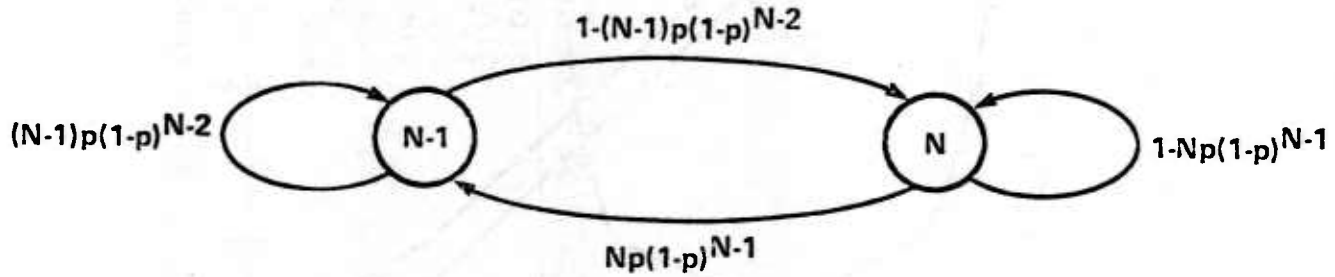

Figure 6 Markov Chain at Saturation ($\lambda = 1$).

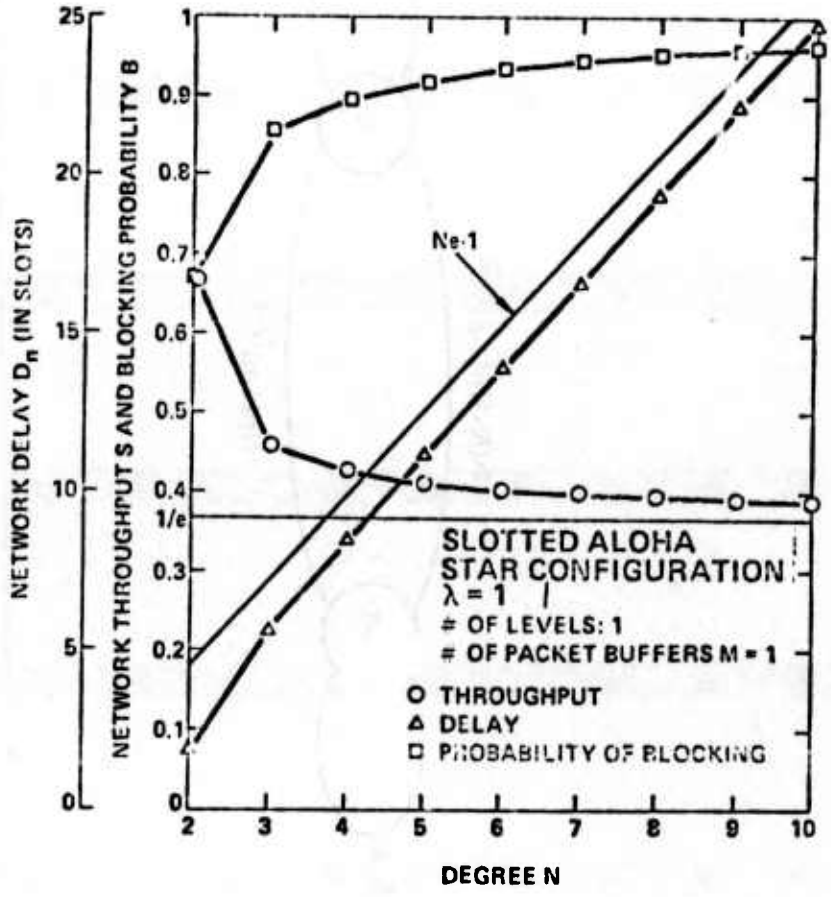

Figure 7 Slotted ALOHA Star Configuration: Performance at Saturation ($\lambda = 1$).

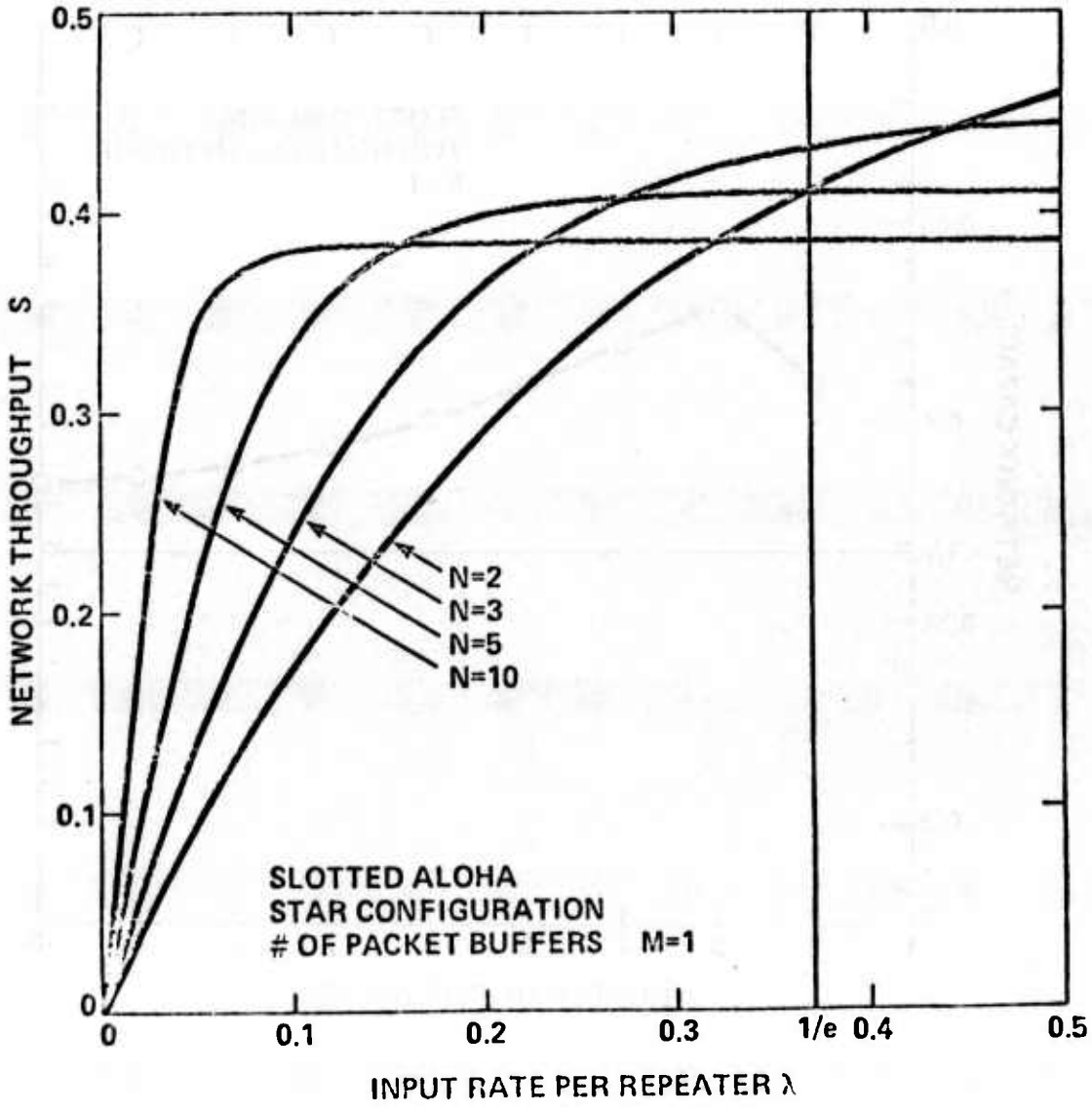

Figure 8 Slotted ALOHA Star Configuration: Throughput versus λ.

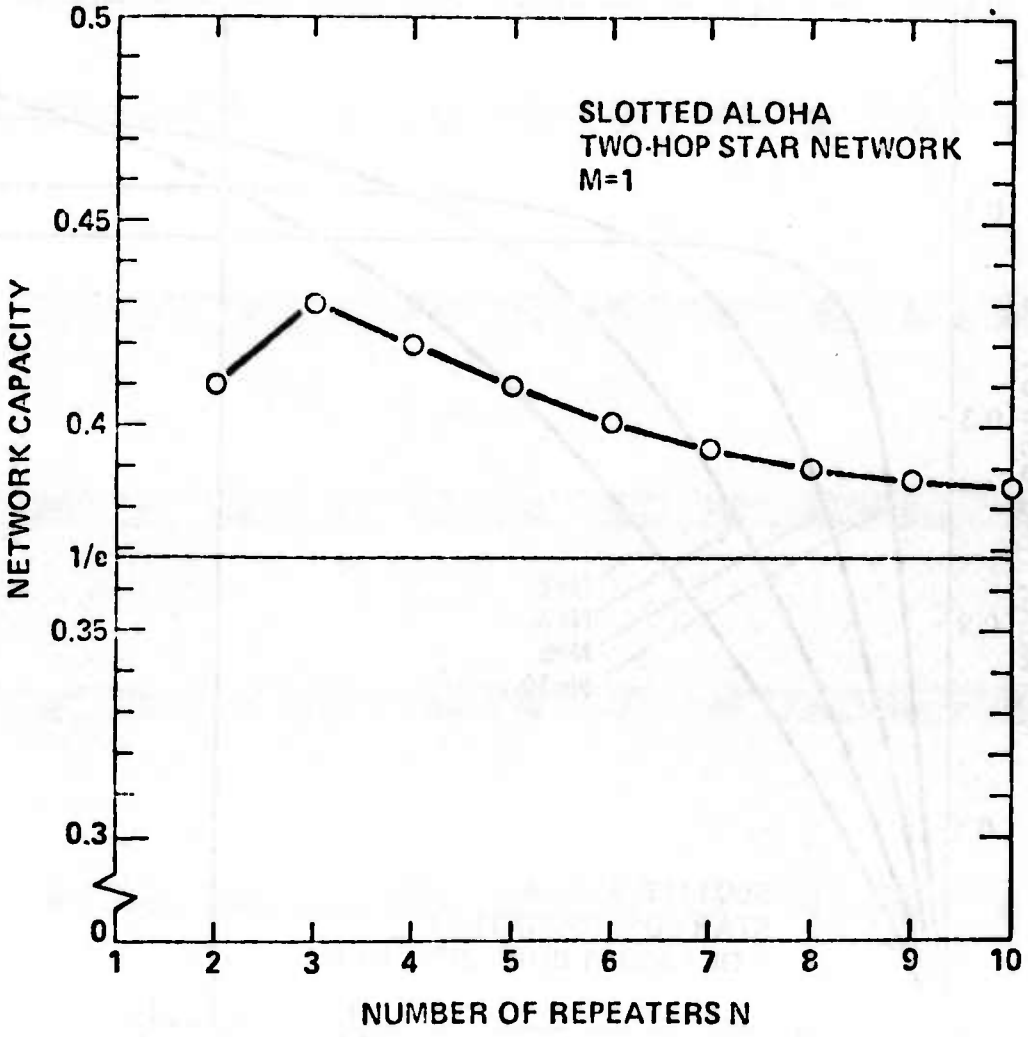

Figure 9 Slotted ALOHA Two-Hop Star Network: Network Capacity versus N.

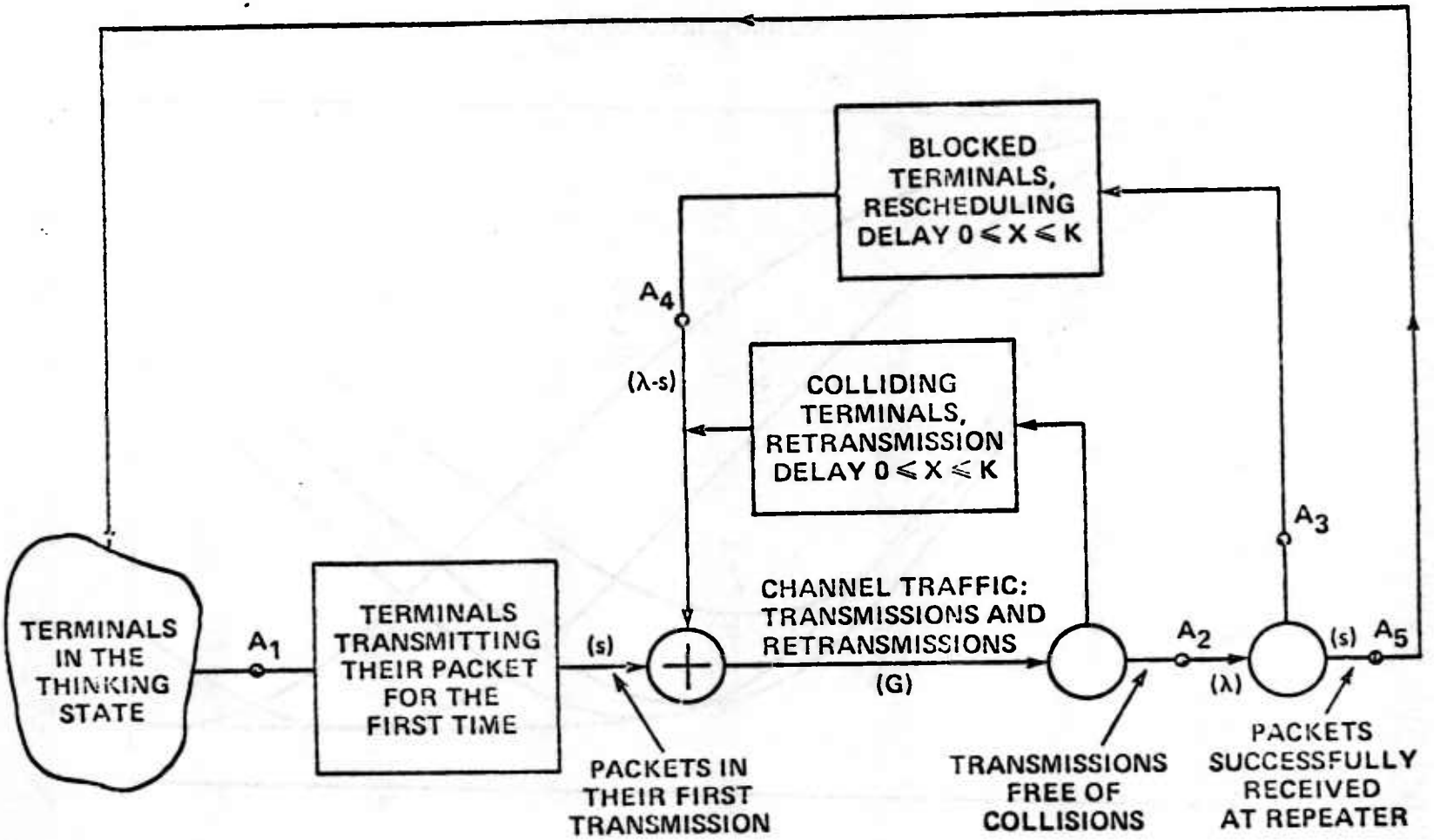

Figure 10 State Diagram for a Population of Terminals.

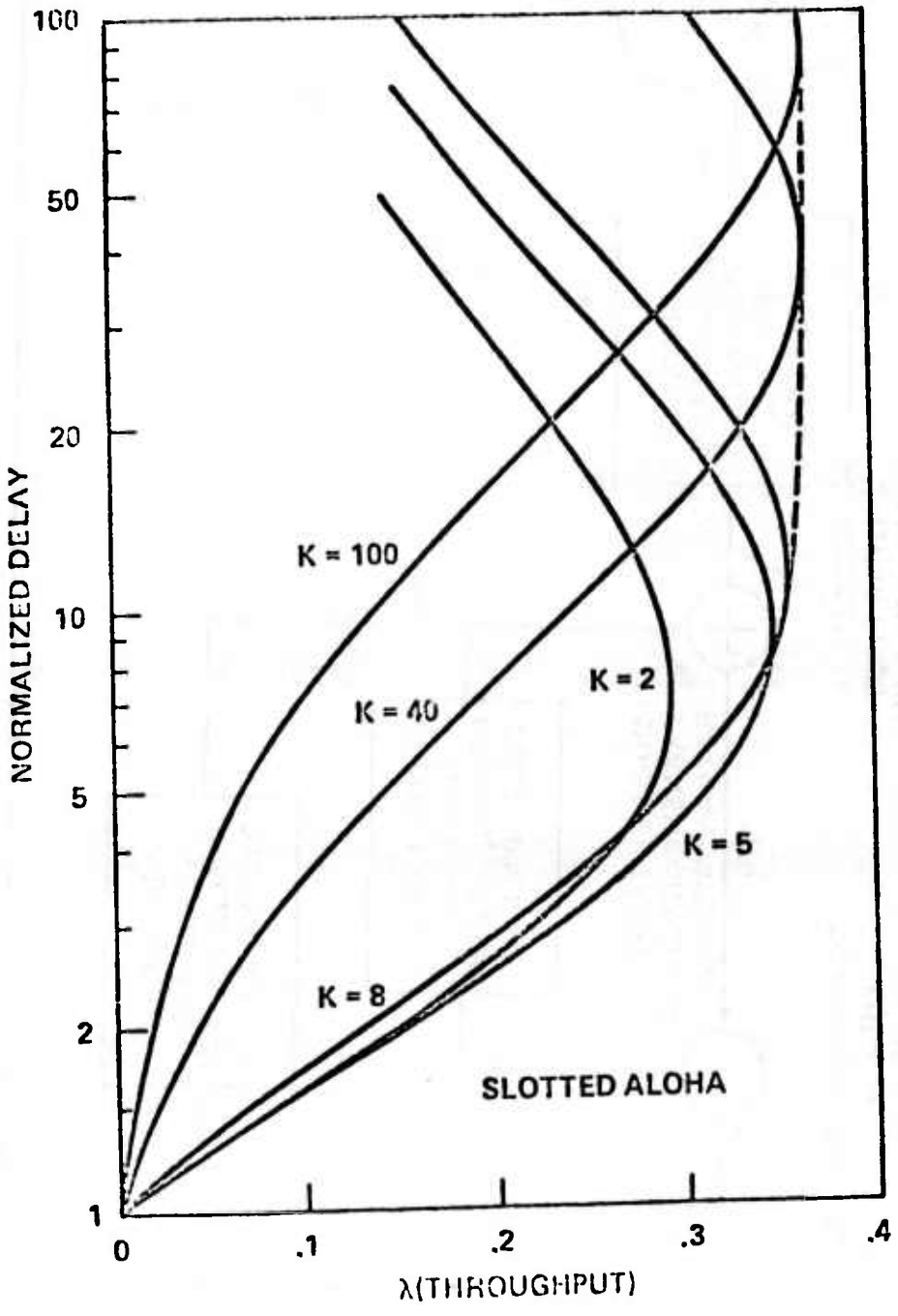

Figure 11 Delay in Slotted ALOHA Channels with Infinite Population.

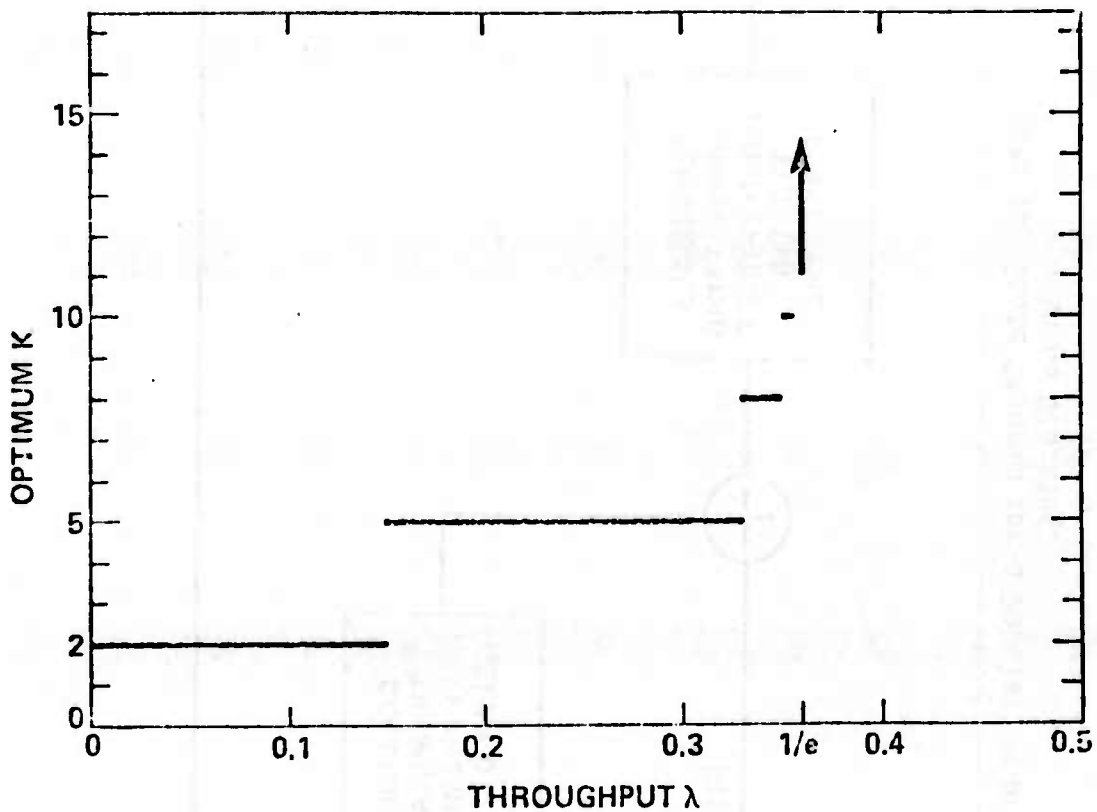

Figure 12 Optimum K versus Throughput

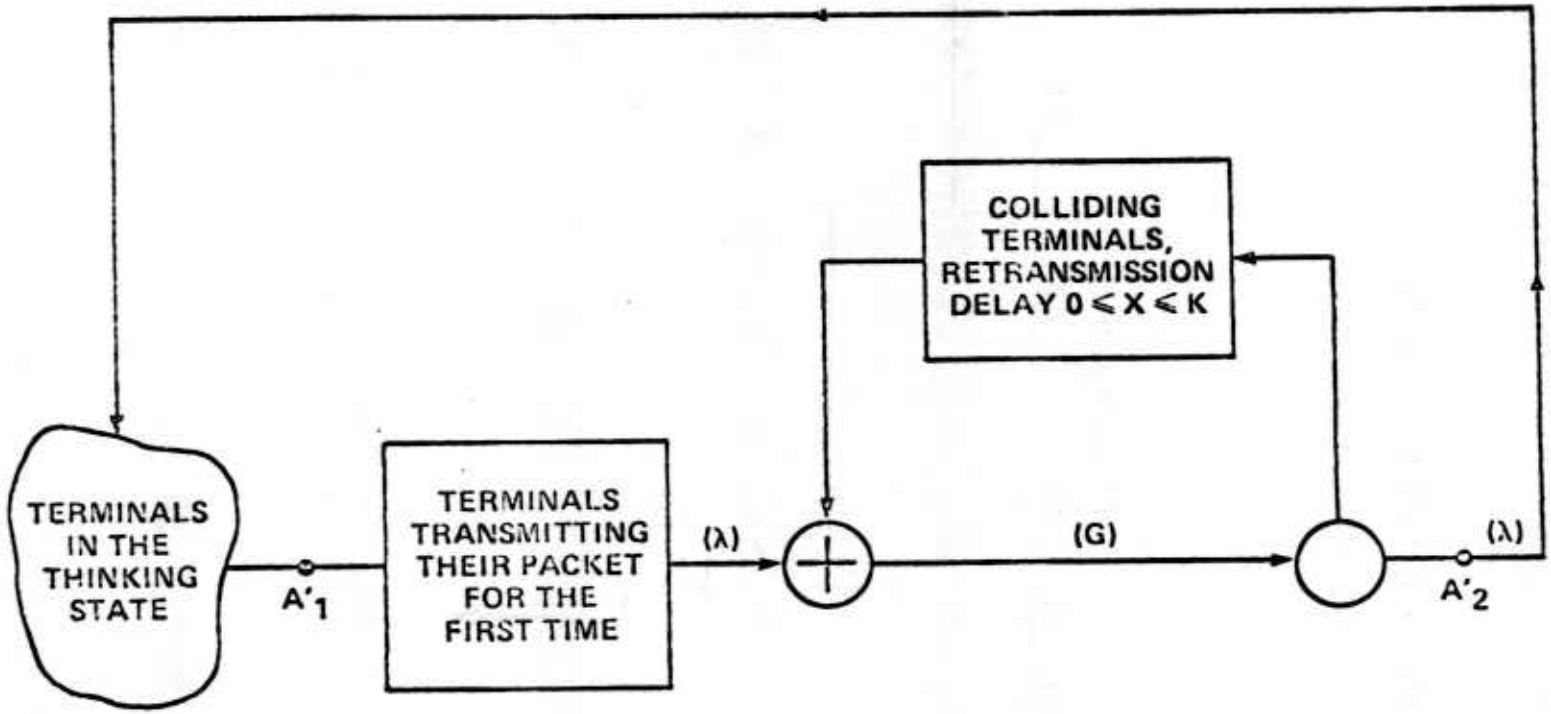

Figure 13 State Diagram for a Population of Terminals with no Blocking.

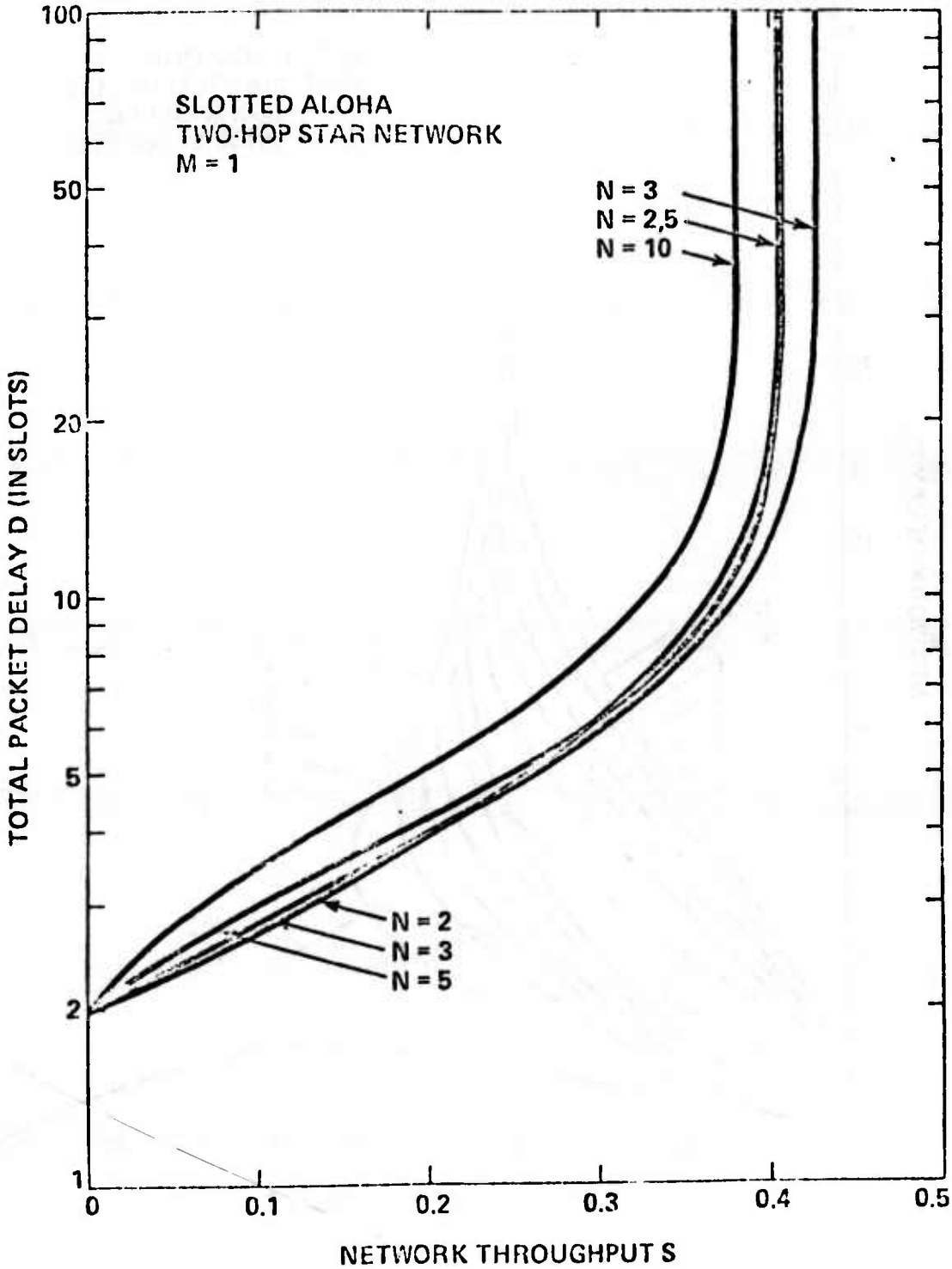

Figure 14 Throughput-Delay Tradeoff in Two-Hop Slotted ALOHA Star Networks

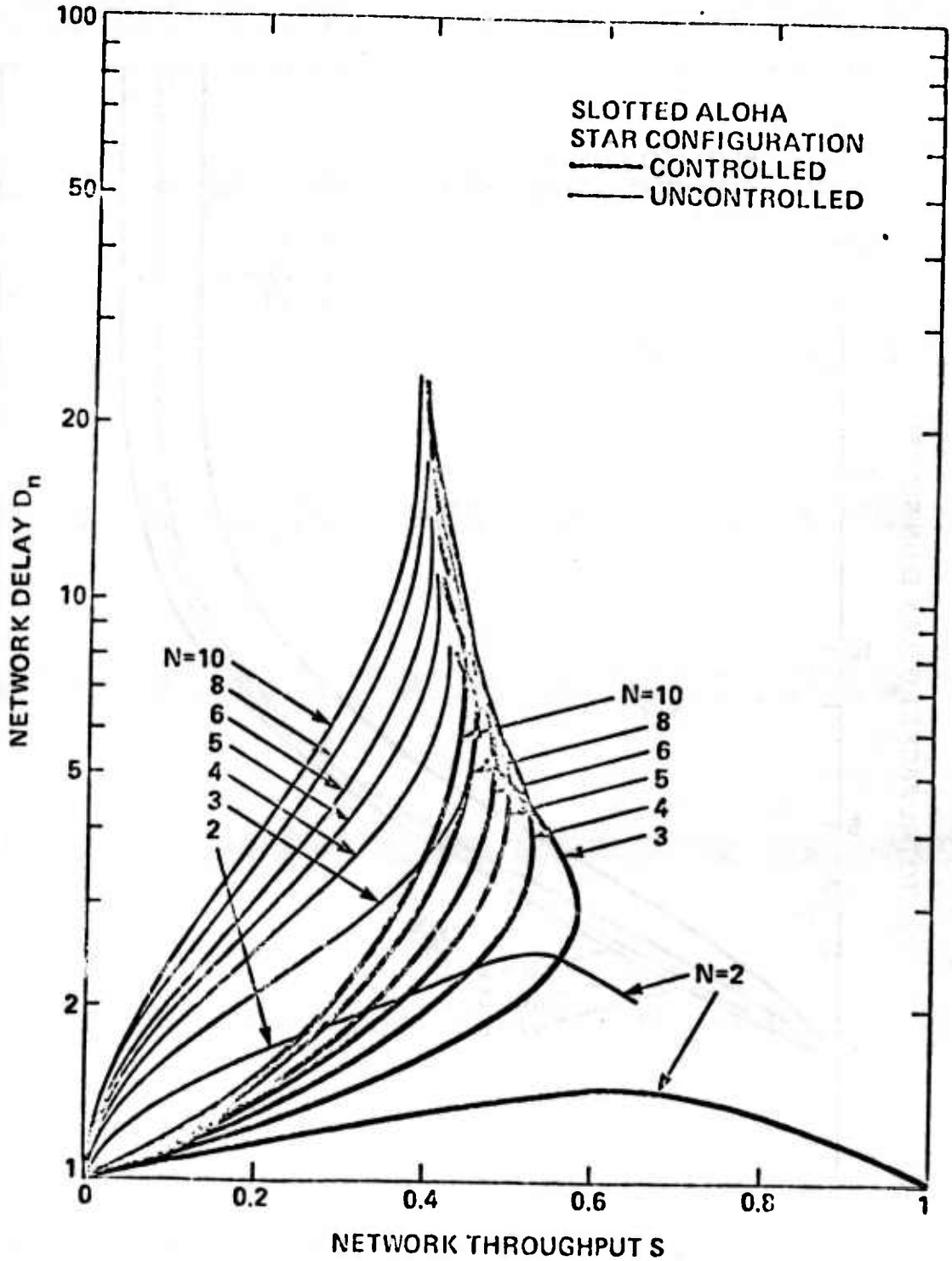

Figure 15 Slotted ALOHA Star Configuration: Optimum (Network) Throughput-Delay Curves for Controlled and Uncontrolled Systems.

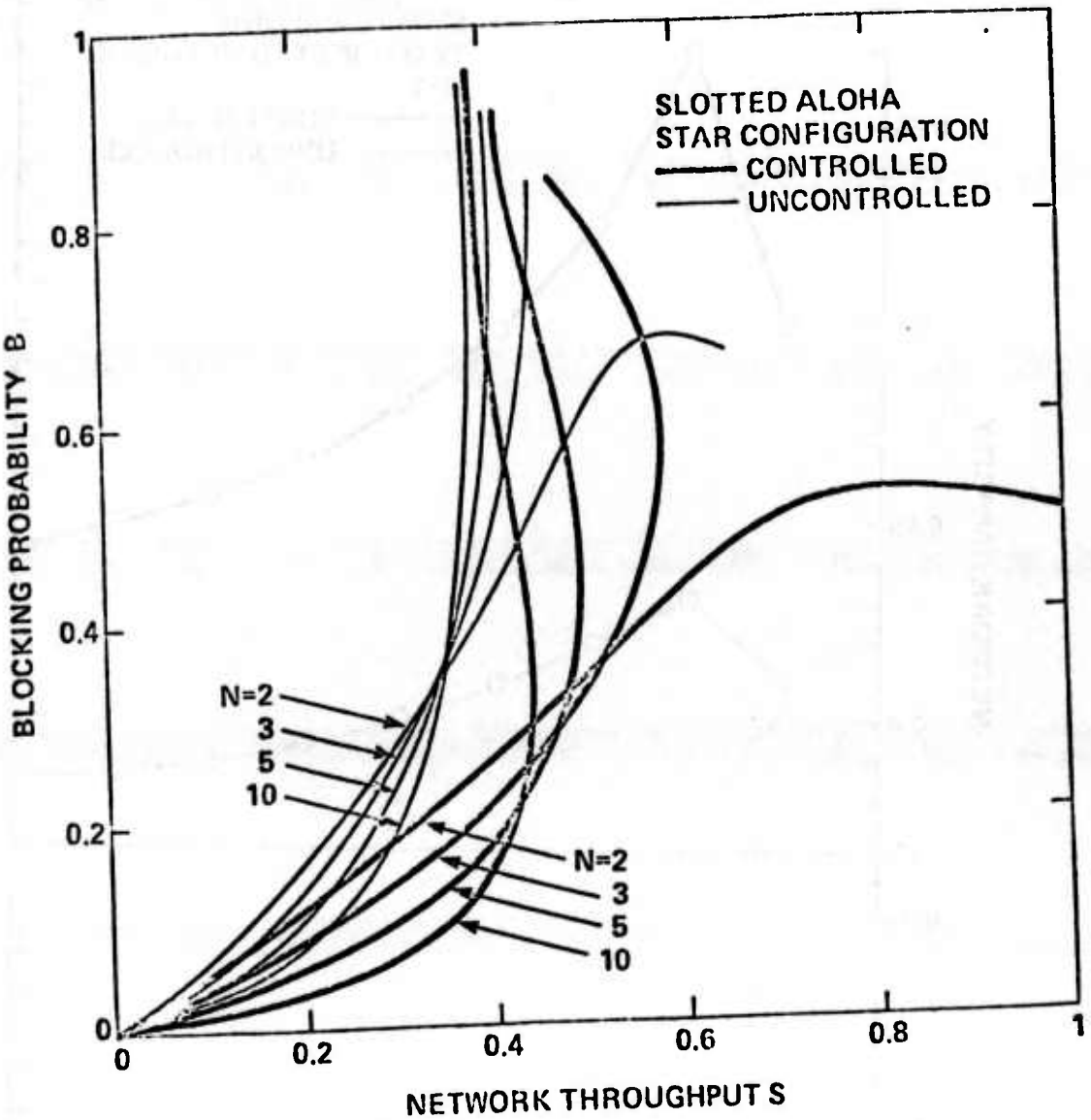

Figure 16 Slotted ALOHA Star Configuration: Minimum Blocking versus S for Controlled and Uncontrolled Systems.

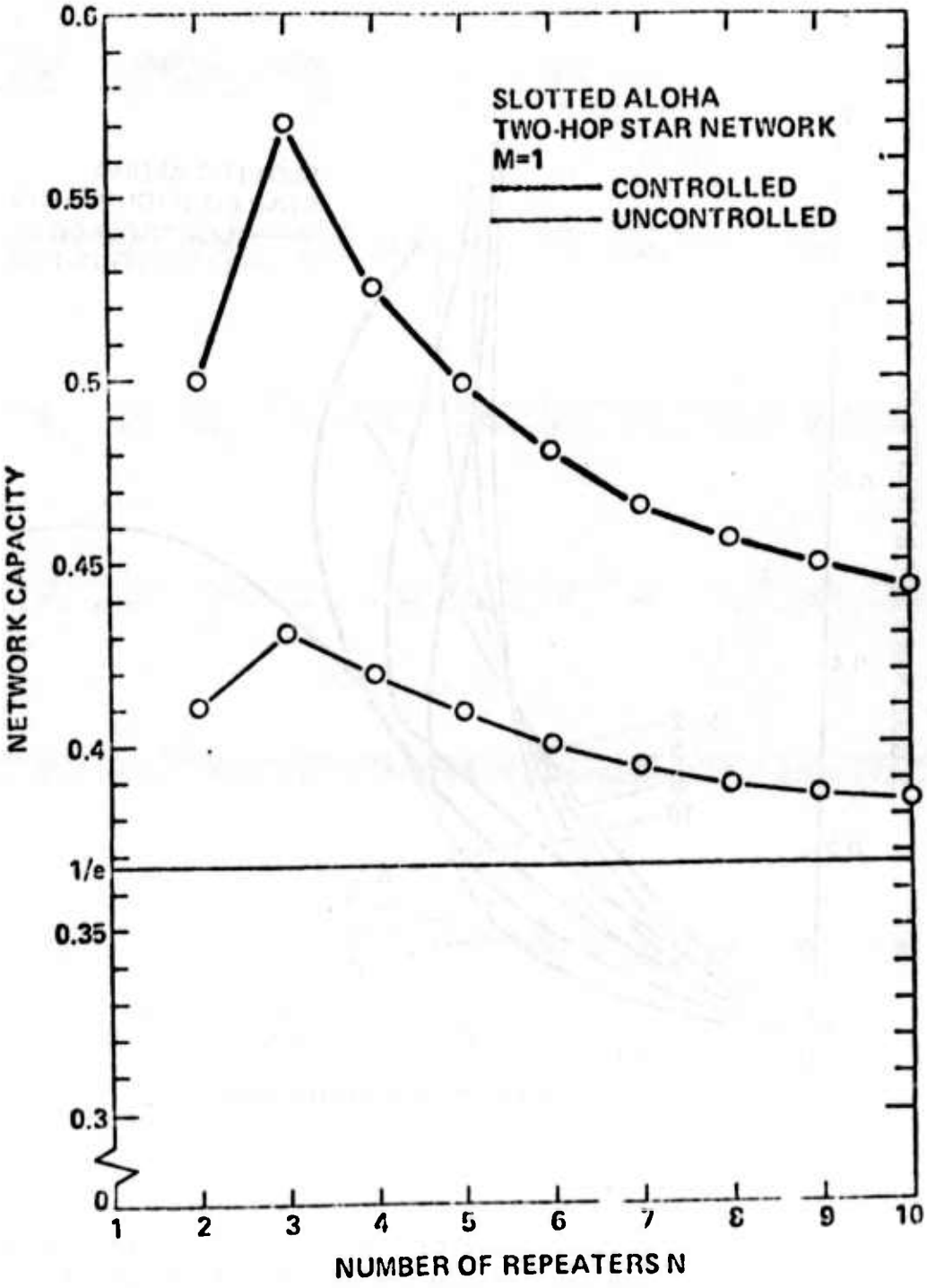

Figure 17 Slotted ALOHA Two-Hop Star Network: Network Capacity versus N.

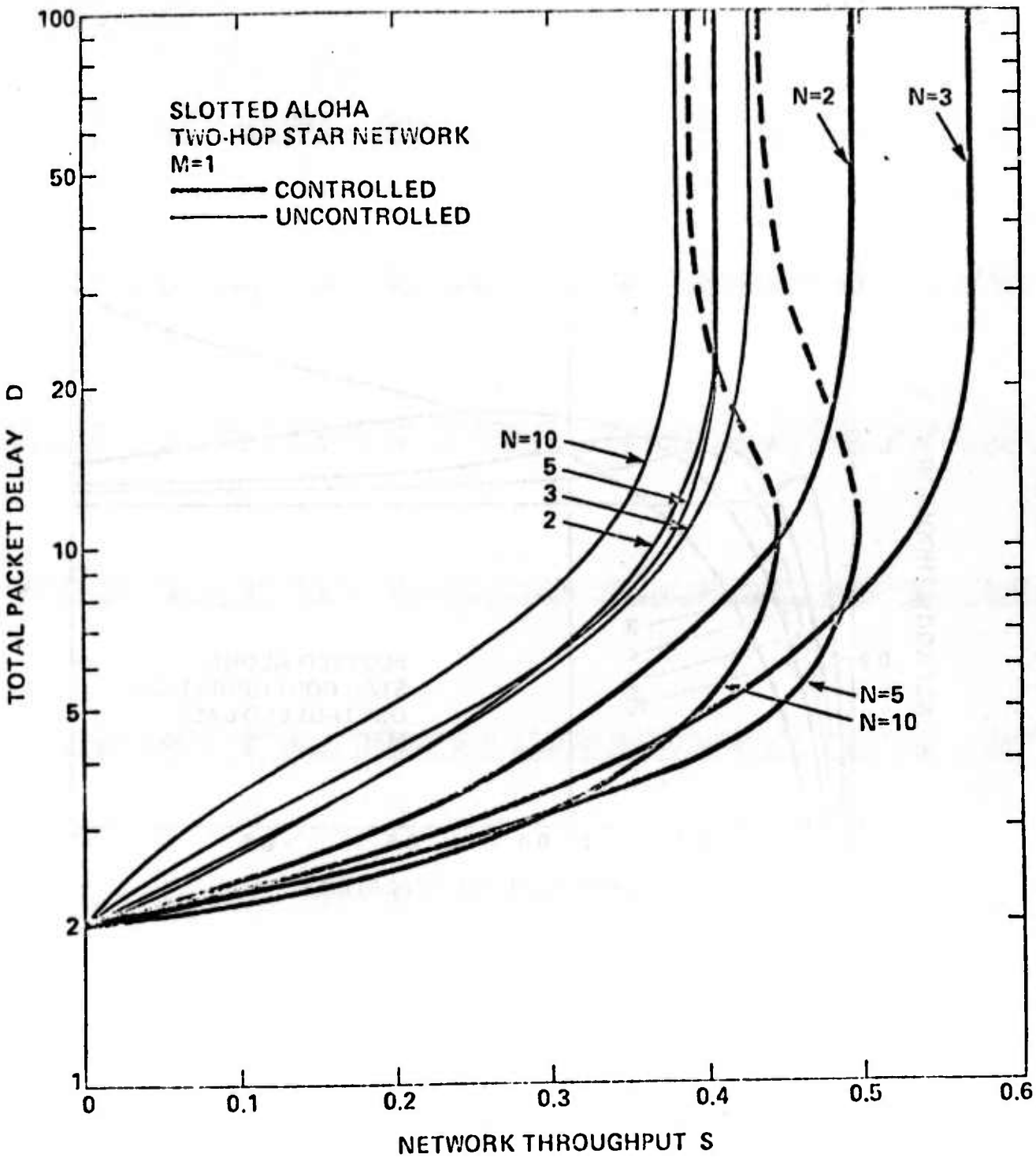

Figure 18 Throughput-Delay Tradeoff in Controlled Two-Hop Slotted ALOHA Star Networks.

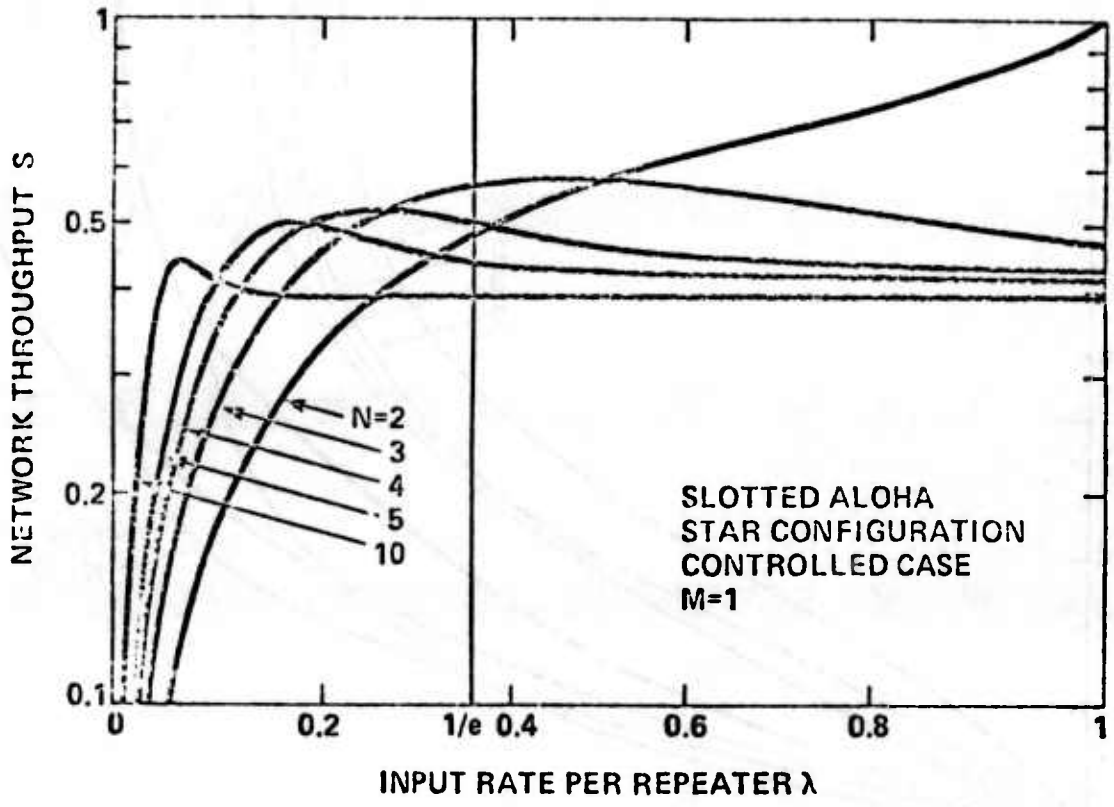

Figure 19 Controlled Slotted ALOHA Star Configurations: Throughput versus λ.

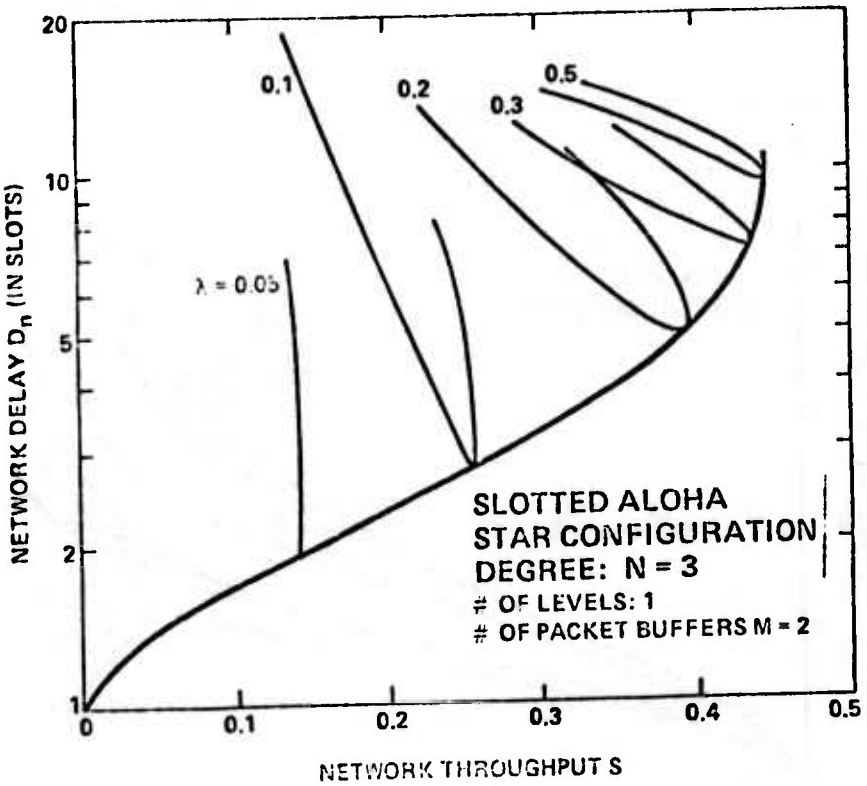

Figure 20 Slotted ALOHA Star Configuration: Delay versus Throughput with M > 1.

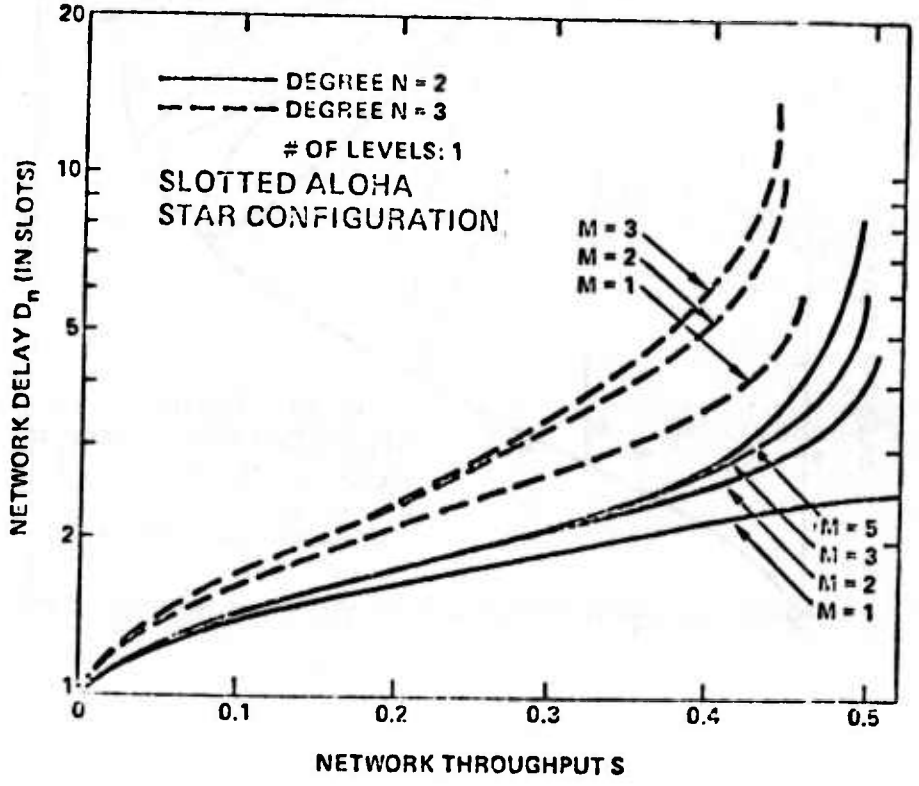

Figure 21 Slotted ALOHA Star Configuration: Minimum Delay versus S for Various Values of M.

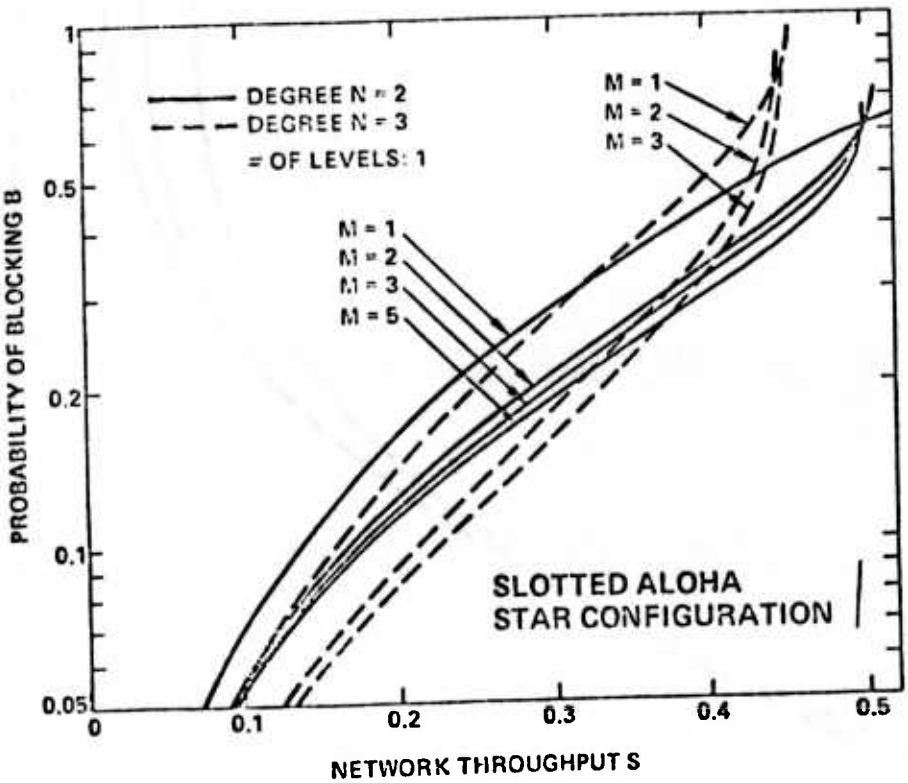

Figure 22 Slotted ALOHA Star Configuration: Minimum Blocking versus S for Various Values of M.

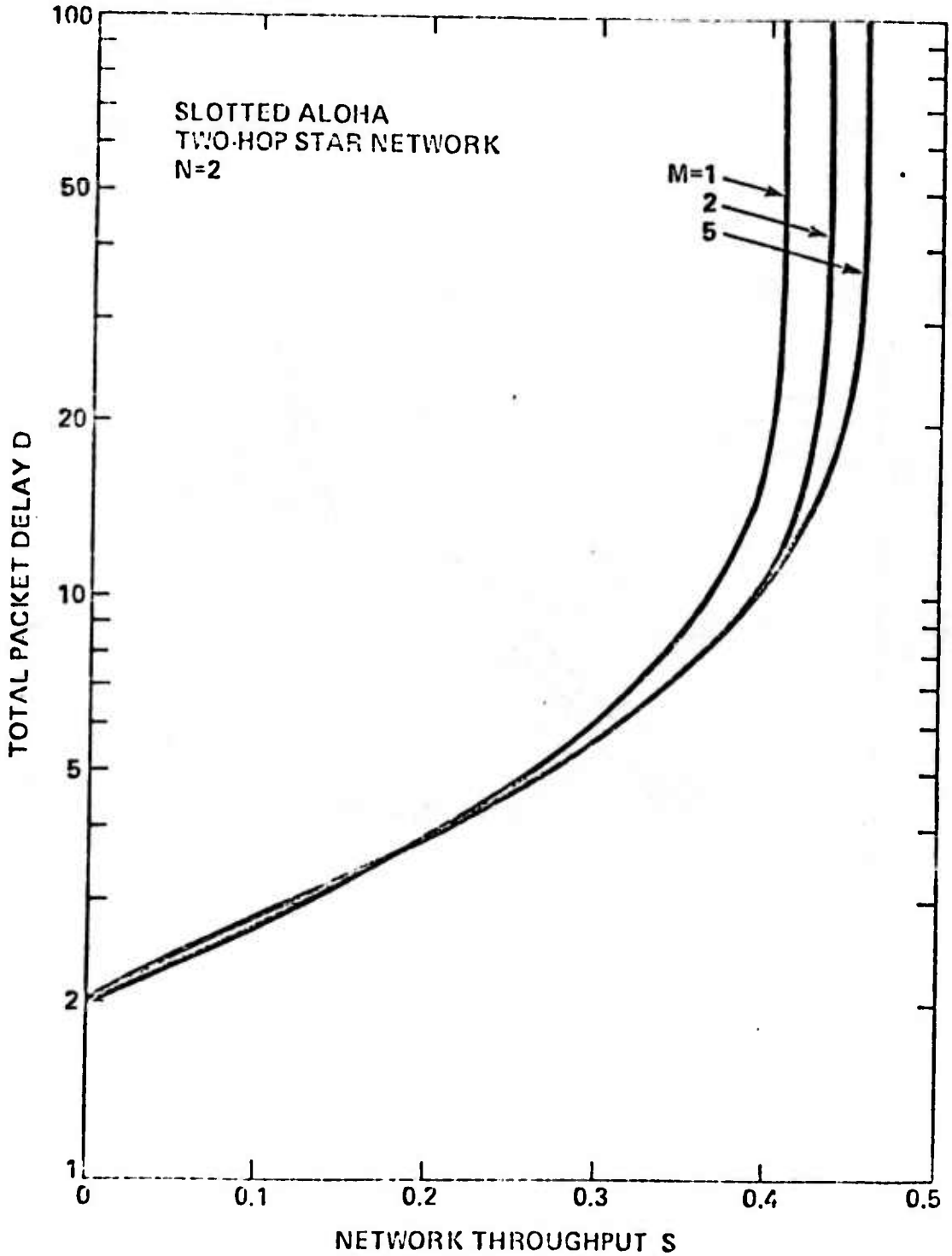

Figure 23 Two-Hop Slotted ALOHA Star Networks: Total Packet Delay versus S for N = 2 and M $\geq$ 1.

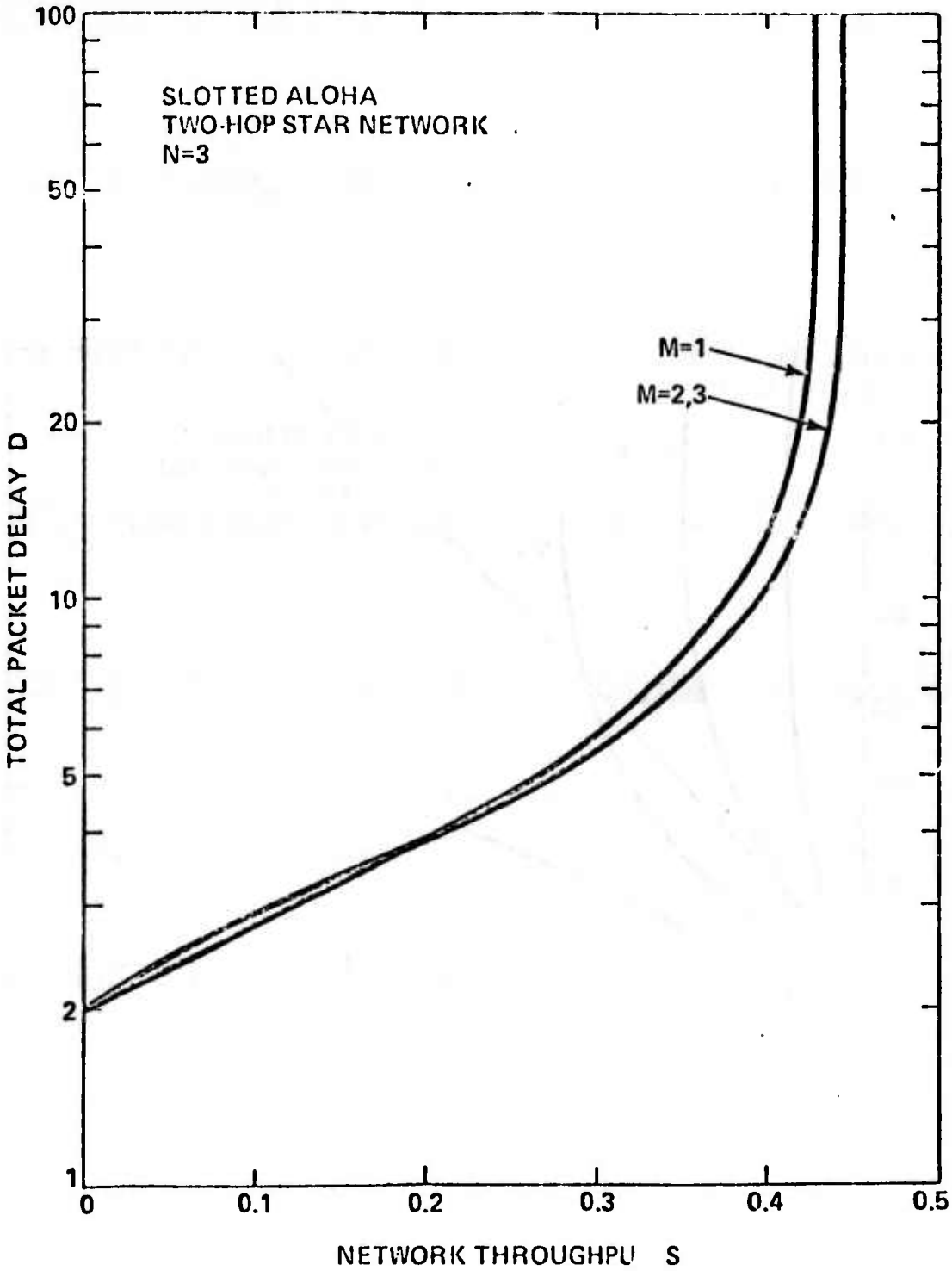

Figure 24 Two-Hop Slotted ALOHA Star Networks: Total Packet Delay versus S for N = 3 and M $\geq$ 1.

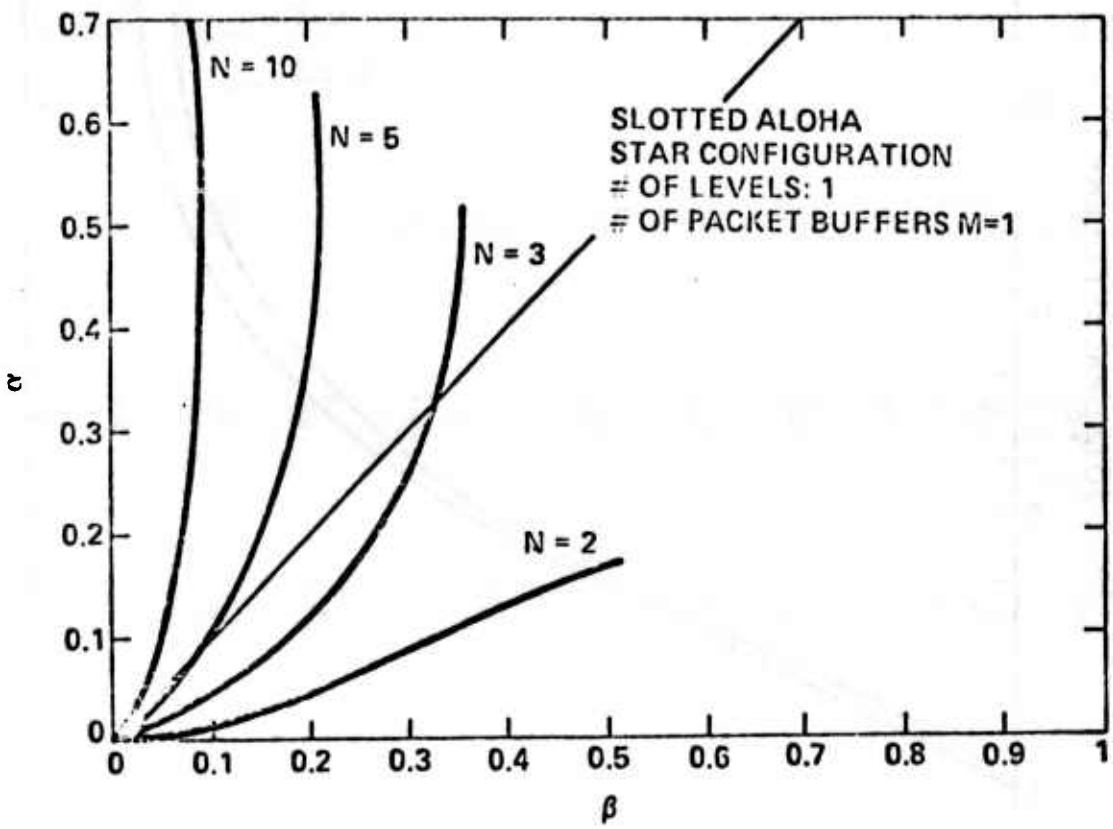

Figure 25 Slotted ALOHA Star Configuration: α versus β at Minimum Blocking.